

From: Swope, Sheridan
Sent: Thursday, June 09, 2005 6:04 PM
To: STIC-Biotech/ChemLib
Subject: 10023515

For 10023515, pls search and interference search:

SID 1 against the NT and AA data bases.

SID 2 against the NT and AA data bases.

SID 3 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

Toby
Port

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 08:57:45 ; Search time 89 Seconds

(without alignments)
3342.899 Million cell updates/sec

Title: US-10-023-515-2
Perfect score: 3079
Sequence: 1 MPQGLTSSASQWCFLLIQP.....PLSLTFLSLIQPPFFCAP 581

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trernbl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2915	94.7	575 2	Q6NTJ32
2	2606	84.6	525 2	Q96DN9
3	2454	79.7	469 2	Q8NRC8
4	2250	73.1	575 2	Q6AM47
5	2094.5	68.0	575 2	Q6AM46
6	1985	64.5	545 2	Q81034
7	1676.5	54.4	361 2	Q95KH3
8	1298.5	42.2	557 1	SASB_ANAPL
9	1282.5	41.7	561 2	Q91W60
10	1281	41.6	559 2	Q8BK48
11	1276.5	41.5	559 2	Q35533
12	1271	41.3	559 2	Q8BM97
13	1266.5	41.1	556 2	Q6PDB7
14	1263.5	41.0	558 2	Q6GMS4
15	1262.5	41.0	561 2	Q70177
16	1260.5	40.9	560 2	Q8R097
17	1257.5	40.8	554 2	Q35535
18	1239.5	40.3	562 2	Q6P306
19	1238	40.2	558 2	Q8K3R0
20	1233.5	40.0	558 2	Q8GZ83
21	1231.5	40.0	532 1	EST2_RABIT
22	1230	39.9	559 1	EST2_HUMAN
23	1213.5	39.4	561 2	Q70631
24	1213.5	39.4	566 2	Q70631
25	1213.5	39.4	566 2	Q86489
26	1210.5	39.3	557 2	Q640T6
27	1199	38.9	561 1	EST1_MESAU
28	1191	38.7	586 2	Q61PK9
29	1187.5	38.6	568 2	Q6PG74
30	1187.5	38.6	572 1	Q810S9
31	1186.5	38.5	554 1	ESTM_MOUSE

32	1183	38.4	565 2	Q95N05	Q95N05 canis fam1
33	1179.5	38.3	571 2	Q6UW8	Q6UW8 homo sapien
34	1179	38.3	568 2	Q726J1	Q726J1 homo sapien
35	1175	38.2	565 2	Q8RD29	Q8RD29 homo sapien
36	1171.5	38.0	568 2	Q8YCU1	Q8YCU1 mus musculu
37	1169.5	38.0	561 2	Q91WU0	Q91WU0 mus musculu
38	1164	37.8	565 2	Q97582	Q97582 sus scrofa
39	1162.5	37.8	555 2	Q6GMJ1	Q6GMJ1 brachydanto
40	1162.5	37.8	566 1	EST1_PIG	Q29550 sus scrofa
41	1159.5	37.7	566 2	Q464Z1	Q464Z1 macaca fasc
42	1150.5	37.4	565 1	EST1_RABIT	P12337 oryctolagus
43	1150	37.3	554 1	ESTN_MOUSE	P23953 mus musculu
44	1149	37.3	565 2	Q35534	Q35534 mesocricetu
45	1144	37.2	549 1	EST1_RAT	P10959 rattus norv

ALIGNMENTS

RESULT 1

ID Q6NTJ32 PRELIMINARY; PRT: 575 AA.

AC Q6NTJ32; 05-JUL-2004 (TRENBLrel. 27, Created)

DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)

DE 05-JUL-2004 (TRENBLrel. 27, Last annotation update)

GN FLJ131547 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RP [1]

RC TISSUE=PCR rescued clones;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiy S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,

RT Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RM [2]

RP SEQUENCE FROM N.A.

RC TISSUE=PCR rescued clones;

RA Director MGC Project;

RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.

DR EMBL; BC069501; AAH69501.1; .

DR HSSP; P12337; IKAY.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR InterPro; IPR002018; Carboxylesterase.

DR InterPro; IPR003379; Ser. catrs.

DR Pfam; PF00135; Coesterase; 1.

DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.

DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.

KW Hydrolase.

SQ SEQUENCE 575 AA; 63926 MW; 6F5B735BDFC9C09 CRC64;

Query Match 94.7%; Score 2915; DB 2; Length 575;
 Best Local Similarity 97.0%; Pred. No. 6.2e-214;
 Matches 553; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 12 WCFPLIQPLIGHRQWCKTSPSAGPQNRRLGNIQKQVTVLGSPPVAVVFLGVPFPAAP 71
 DB 14 MAIWLVLAP-----TKGPSAGPQRNRLGNIQKQVTVLGSPPVAVVFLGVPFPAAP 65
 QY 72 PLGSLRTNPQAPSPMNLREATSYRPLCTQNSEMLLDQMLKVHPKFGVSEDCLYIN 131
 DB 66 PLGSLRTNPQAPSPMNLREATSYRPLCTQNSEMLLDQMLKVHPKFGVSEDCLYIN 125
 QY 132 IYAPAHADTGSKLPLVLPVFPFGAFKTSASIFDGSALAAVEDLVVVVQVRLGIFGFFTT 191
 DB 126 IYAPAHADTGSKLPLVLPVFPFGAFKTSASIFDGSALAAVEDLVVVVQVRLGIFGFFTT 185
 QY 192 WDQHAAGNNAFKQVVALSVQKNIEPFGDPSSVTIIFGSSAGAISVSSLLISPMKGLF 251
 DB 186 WDQHAAGNNAFKQVVALSVQKNIEPFGDPSSVTIIFGSSAGAISVSSLLISPMKGLF 245
 QY 252 HKAIMSGVAIIPIYLAHDYKSEDLQVVAHPCGNNAASEALLRCLRTKPSKELLTLSQ 311
 DB 246 HKAIMSGVAIIPIYLAHDYKSEDLQVVAHPCGNNAASEALLRCLRTKPSKELLTLSQ 305
 QY 312 KTKSFTFRVVDGAFEPNEPDLISQKAFKAIPSIIGVNNHECGFLIPMKEAPFLLSGSNKS 371
 DB 306 KTKSFTFRVVDGAFEPNEPDLISQKAFKAIPSIIGVNNHECGFLIPMKEAPFLLSGSNKS 365
 QY 372 LALHLIQLNLIHPQVLIHVAHEYPHDKSLTEIRPSLLDLIDGVFFVVPALITARYRD 431
 DB 366 LALHLIQLNLIHPQVLIHVAHEYPHDKSLTEIRPSLLDLIDGVFFVVPALITARYRD 425
 QY 432 AGAPVYFEFRHRPOCFEDTKPAFYKADHADVRFVFGGAFPLKGIIVMEGATEEKLLS 491
 DB 426 AGAPVYFEFRHRPOCFEDTKPAFYKADHADVRFVFGGAFPLKGIIVMEGATEEKLLS 485
 QY 492 RRMCMYATPATGPNPNDLSLMPAYNLTEQYLOLDINMSIGRLKSPRYDFTSTIPL 551
 DB 486 RRMCMYATPATGPNPNDLSLMPAYNLTEQYLOLDINMSIGRLKSPRYDFTSTIPL 545
 QY 552 ILSASDMLHSPLSLTFSLLOPFFFCAP 581
 DB 546 ILSASDMLHSPLSLTFSLLOPFFFCAP 575

RESULT 2
 Q96DN9 PRELIMINARY; PRT; 525 AA.
 AC Q96DN9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 25-OCT-2004 (Tremblrel. 19, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE Hypothetical protein FLJ31547.
 GN Name=FLJ31547;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Euhetia; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RA [1]
 RA SEQUENCE FROM N.A.
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Odayashi M., Nish T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M., Shiratori A.,
 RA Sudo H., Hosioki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y.,
 RA Ninomiya K., Itohishi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimura M., Watanabe M., Hirooka S., Chiba Y., Iehida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Houchi T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,

RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotaka S.,
 RA Yorioka Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetake K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki S.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura Y., Kikuchi H., Maeno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Iguchi T., Sugano S.,
 RA "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs".
 RL Nat. Genet. 36:40-45 (2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PCR rescued clones;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Udell T.B., Toehlyuk S., Carminci P., Prange C.,
 RA Baha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Roark S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Keteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzyzinski M.I., Skalka U., Smalins D.E., Scherach A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PCR rescued clones;
 RA Straubeberg R.;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL: AK056109; BAB71094.1; -;
 DR EMBL: BC069548; AAH69548.1; -;
 DR HSSP: P12337; 1k4Y.
 DR GO: GO:0016787; F:Hydrolase activity; IEA.
 DR InterPro: IPR002018; Carboxylesterase.
 DR InterPro: IPR000379; Ser esters.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYL-ESTERASE B_1; 1.
 DR PROSITE: PS00941; CARBOXYL-ESTERASE B_2; 1.
 KW Hydrolase, Hypothetical protein.
 SQ SEQUENCE 525 AA; 58200 MW; 7724878B8A22F215 CRC64;

Query Match 84.6%; Score 2606; DB 2; Length 525;
 Best Local Similarity 90.9%; Pred. No. 2.2e-190;
 Matches 501; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

QY 31 GPSAGPQRNRLGNIQKQVTVLGSPPVAVVFLGVPFPAAPPLGSLFTNPQAPSPMDNL 90
 DB 25 GPSAGPQRNRLGNIQKQVTVLGSPPVAVVFLGVPFPAAPPLGSLFTNPQAPSPMDNL 84
 QY 91 REATSPMLCTQNSEMLLDQMLKVHPKFGVSEDCLYINITYAPAHADTGSKLPLVLPF 150
 DB 85 REATSPMLCTQNSEMLLDQMLKVHPKFGVSEDCLYINITYAPAHADTGSKLPLVLPF 144

```

QY 151 PGGAFTKGSASIFDGSALAAVEDLVVVVQYRLGIFGFTTMDQHPAGNMAFKQVAAVS 210
DB 145 PGGAFTKGSASIFDGSALAAVEDLVVVVQYRLGIFGFTTMDQHPAGNMAFKQVAAVS 204
QY 211 WVQKNIEFFGGDPSVITFGESAGAISSVLSLSPMAGLPHKAIMESGVAIIPYLEAHD 270
DB 205 WVQKNIEFFGGDPSVITFGESAGAISSVLSLSPMAGLPHKAIMESGVAIIPYLEAHD 264
QY 271 YEKESDQVVAHPCGNNADEBALRCRTKPSKELLTSSQKTSFTFVVGAFPEPNEPL 330
DB 265 YEKESDQVVAHPCGNNADEBALRCRTKPSKELLTSSQKTSFTFVVGAFPEPNEPL 324
QY 331 DLSQKAFKAIPIIIGVNNHRCGFLPMKKAPELISGNSKSLALHIONLHIIPQYIHL 390
DB 325 DLSQKAFKAIPIIIGVNNHRCGFLPMKKAPELISGNSKSLALHIONLHIIPQYIHL 384
QY 391 VANEYFHDKSLTEIRDSLLDLGDVFPVVPALITARYHDAQAVYIEFRHPOCED 450
DB 385 VANEYFHDKSLTEIRDSLLDLGDVFPVVPALITARYHDAQAVYIEFRHPOCED 424
QY 451 TKPAFVADHADVRFVFGCAFVKGDIMFEGATEEERKLSRKMKKWATPARTGNPNCN 510
DB 425 -----EGATEEERKLSRKMKKWATPARTGNPNCN 454
QY 511 DLSMPAYNLTEQYQLDLNMSLQRLKEPRVDFWTSTIPLISASDMLHSPSLSTPLS 570
DB 455 DLSMPAYNLTEQYQLDLNMSLQRLKEPRVDFWTSTIPLISASDMLHSPSLSTPLS 514
QY 571 LLOPFFFCAP 581
DB 515 LLOPFFFCAP 525

```

RESULT 3

```

Q8N8C8 PRELIMINARY; PRT; 469 AA.
AC Q8N8C8;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ33678.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y.,
RA Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki M.,
RA Tanai H., Kimura M., Watanabe M., Hirano K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuki H., Oshima H., Sasaki N., Aochi S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Teraishi Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe K., Kumagai A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fujiwara T.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,

```

```

RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuh Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AK090997; BAC03565.1; -.
DR HSP: P12337; IKAY.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser_esterase.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
DR KW Hydrolyase.
SQ SEQUENCE 469 AA; 52312 MW; 7A188F68C10A9080 CRC64;

```

Query Match 79.7%; Score 2454; DB 2; Length 469;
 Best Local Similarity 99.6%; Pred. No. 7, 8e-179; Mismatches 0; Indels 0; Gaps 0;
 Matches 467; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 113 MLKVHYPRFVSEDCLYNITVYAPAHADGSKLPVLVWPPGAFKTSIFDGSALAAVE 172
DB 1 MLKVHYPRFVSEDCLYNITVYAPAHADGSKLPVLVWPPGAFKTSIFDGSALAAVE 60
QY 173 DVLVVVQYRLGIFGFTTMDQHPAGNMAFKQVAAVSQKNIEFFGGDPSVITFGES 232
DB 61 DVLVVVQYRLGIFGFTTMDQHPAGNMAFKQVAAVSQKNIEFFGGDPSVITFGES 120
QY 223 AGAIVSSSLISPMKAGLPHKAIMESGVAIIPYLEAHDYKESDQVVAHPCGNNADE 292
DB 121 AGAIVSSSLISPMKAGLPHKAIMESGVAIIPYLEAHDYKESDQVVAHPCGNNADE 180
QY 293 ALLRCLRTKPSKELLTSSQKTSFTFVVGAFPEPNEPLDLSQKAFKAIPIIIGVNNHC 352
DB 181 ALLRCLRTKPSKELLTSSQKTSFTFVVGAFPEPNEPLDLSQKAFKAIPIIIGVNNHC 240
QY 353 GFLPMKKAPELISGNSKSLALHIONLHIIPQYIHLVANEYFHDKSLTEIRDSLLD 412
DB 241 GFLPMKKAPELISGNSKSLALHIONLHIIPQYIHLVANEYFHDKSLTEIRDSLLD 300
QY 413 LGDVFVVPALITARYHDAQAVYIEFRHPOCEDTKPAFVADHADVRFVFGCAF 472
DB 301 LGDVFVVPALITARYHDAQAVYIEFRHPOCEDTKPAFVADHADVRFVFGCAF 360
QY 473 LKGDIVMEGATEEERKLSRKMKKWATPARTGNPNCNLSMPAYNLTEQYQLDLNMS 532
DB 361 LKGDIVMEGATEEERKLSRKMKKWATPARTGNPNCNLSMPAYNLTEQYQLDLNMS 420
QY 533 LGQRIKEPRVDFWTSTIPLISASDMLHSPSLSTPLSLOPFFFCAP 581
DB 421 LGQRIKEPRVDFWTSTIPLISASDMLHSPSLSTPLSLOPFFFCAP 469

```

RESULT 4

```

Q6AW47 PRELIMINARY; PRT; 575 AA.
AC Q6AW47;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxin;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Placentalia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyazaki M., Yamashita T., Taira H., Suzuki A.;
RT "cauxin family protein.";

```

RL Submitted (AUG-2004) to the EMBL/GenBank/DBEJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB166392; BAD5015.1; -.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR000379; Ser_cstrs.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 KW Hydrolase.
 SQ SEQUENCE 575 AA; 63620 MW; 1005C35B82E1183D CRC64;

Query Match	73.1%	Score 2250;	DB 2;	Length 575,
Best Local Similarity	75.7%	Prod No 40123		

	Matches	41 %	Conservative	61 %	Mismatches	73 %	Indels	0 %	Gaps	0 %
QY	31	GPSAEQPORNRRLMIOGQOVTVLGSPPVAVNVFLGVPPAADPLGSLRPTNPQAPMNTL	90							
Dp	25	GPADADAVRSLRCLGWKQKATVAGSTIPVAVFSLIPPAAPLPGLRKRKRPPALLMND	84							
QY	91	REATSYENLCLQNSEWILLDQMLKVHPKRGVSEDCYLIINIYAPAHADTSGSKLPVLMWF	150							
Dp	85	RDATSYEPLCLQNSWILLSDQHLFKVHPNLVESEDCYLIINIYAPAHANTSGSKLPVLMWF	144							
QY	151	PGCAFRTKGSASIFPGSALAAYEDVLVVVVQYRLGIFGFTTWDQHPAGNMAFKQVVALS	210							
Dp	145	PGCAFRTKGSASIFPGSALAAYEDVLVTTQYRLGIFGFTKGDDHAPGNMAFLDQALALT	204							
QY	211	WVQGNIEFFPGGDDPSVTIFGGSAGAIYSSLSLSPMAKGLPHKAIMESGVALIPLYLAHD	270							
Dp	205	WVQGNIEFFPGGDDPSVTIIFGGSAGAIYSSGLVLSPPMAKGLPHKAIMESGVALIIPFLAAPD	264							
QY	271	YKESSEDLOVVAHPFCGNNASDSEALLRCLRTPSKELLTLSQTKSFFRVVDGAFPNBPL	330							
Dp	265	DERNEDLOVIAIRLQCGCNVSDVALLOCLTRAKSSSELLDINKTKYSFFRVVDGFFPDEPL	324							
QY	331	DLISQCAFKAIPSIIGVNNHCEGFLPMPKAPBILSGNSKSLMHLIONTILHPIPOYLAL	390							
Dp	325	DLITKXTFNSIIPSVIGVNNHCEGFLPMPKEPBIILSGNSKSLMHLIHRVLHINPOLYTL	384							
QY	391	VANVEYHDKSLRLEIRBDSLLDLDGVFVFPVLPALITAIYHHDAGAPVYFYFRRHPQCFED	450							
Dp	385	VADQYFYNNKHSPEVEIRBDSFLDLGDVFFVVPVGYTAIYHRDAGAPVYFYFRRHPQCFED	444							
QY	451	TKPAFFKADHADDEVRFVFGGAFPLKGDIVMEEGATTEEBKLSRRKMKYMAFFATGPDNGN	510							
Dp	445	TKPAFFKADHSDIRVFPFGGAFPLKGNINWFEGATTEEBKLSRRKMKYMAWAFATGPDNGR	504							
QY	511	DLSLWPAVNLTEQYQLDNLMSLIGQRLKEBRVDFWTSTIPLILASDMLHSPLSSTFSS	570							
Dp	505	GLPLWPAVYQSSEQYKLTDLNISTVGQRLKQGEVDFWSDTILPLIMSMSTAPRGPVPLLSLS	564							
QY	571	LLQPPFFFCAP	581							
Dp	565	VLLPFLFESSAP	575							

RT "cauxit"family proteain.":
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family
 DR GO; GO:0016393; BADA5016.1.; -.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR InterPro: IPR002018; CarboxylesteraseB.
 DR InterPro: IPR000379; Ser esters.
 DR Pfam; PF00135; Coesterase; 1.
 DR ProSITE; PS00123; CARBOXYLESTERASE_B_1; 1.
 DR ProSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 KM Hydrolyase.
 SQ SEQUENCE 575 AA; 6416 MW; 3EB9D85981D9DE0A CRC64;

Query Match	Score	DB 2	Length
68.0%;	2094.5;		575;

	Matches	390; Conservative	69; Mismatches	99; Indels	11; Gaps	1
QY	12	WCFEFLIDPLIGHRMWKTGSPASGQPNRRLRHCIGQKQVTVLGSVPNNVFLGVPFPAAP				71
Db	17	WTFGALII-----GSVTPEPRRYTKLGWVOGKAQTVLGRLEPNNVFLGIPFPAAP				65
QY	72	PLGSLRFTNPOPASFPMDLREATSYSPNLCIQNSEWMLLDQHMVKHYRPKFGSEBQCLYLN				131
Db	66	PLGSLRFSKQPPRIPMQDLREATAIYPRPLCQNLQEMFLFYQNLTKSYRPLGNSBQCLYLN				125
QY	132	IYAPHADTGSCLKPYLWFFPGAGFKTGSASIFDGSALAAVEDLVVWVYQYRLGIFGFTT				191
Db	126	IYAPCHANNKSLSPYMWIIPGGGFETGSASIFDGSALAYEDLVVTVIGYRLGIFGFTT				185
QY	192	MDQHAPGNMAFKDQVAAALSMWQKNIERFGGDPBSYVIRFESAGATSVSGLIISPAKCLF				251
Db	186	QNHAPGNMAWMDQALALLWREKIKFGGNGPVSYVIRFNSAGATSISSLIISPLSADLF				245
QY	252	HKAINESGVALIPIYLEADHYEKSEDLQVWVAFQGNNAASBALKRLCRKPKSEKLLTTSQ				311
Db	246	HRAIMQSGVALIIPBLSKSDMDLKHDIQVYANVCDCVNSKALKCLCEKSKSLIEMLSIQ				305
QY	312	KTKSGTRVVDGAFPPNEDLDLISQAKAFKALPSIIIGVNNHCGFLLPMKABEILISGSNKS				371
Db	306	KAKSFTFRVVDGSPFSEBPELLEISQTKLTIKVPISIIIGVNNDECGYILIPVADTPEILLGSNES				365
QY	372	LALHLIQNLIHPIPOYHLVANEYFPHDKISLREIRSLDLIDGVPFVYVPLIARHYRD				431
Db	366	TALTLHTLTLHPIPOHLYIVTKYEFHGKHSPTDIRPTLLDLIDGVPFVYVPLIARHYRD				425
QY	432	AGAPVYFYEFBRPOCFEDTRPAFYADADAEVRFYFGGAFILKQDIYVFEAGATEEBKLLS				491
Db	426	SGGPVYFYEFQRHRCQGNRPAPFVADITDEIRVPFGGPFILKQDVVFEATEEBKLLS				485
QY	492	RKRMKTYATPAKTGPNQNDLSLMPAYNLTEQYIQLDLMSLGQRLKEPRVDFTSTIPL				551
Db	486	RKMKKYANAFARSGDPEAGADLPMPVYDENEGYIELDVNISGRRLKQGRVRFPTDTLPL				545
QY	552	ILTSADMLHSPLSLTFLSLQFFPFCA	560			
Db	546	ILTSASKALSPFTSLILSLISLSPVLLSAA	574			

RESULT 5		
Q6AM46		
ID	Q6AM46	PRELIMINARY;
AC	Q6AM46;	PRT, 575 AA.
DT	25-OCT-2004	(TREMBLrel. 28, Created)
DT	25-OCT-2004	(TREMBLrel. 28, Last sequence update)
DT	25-OCT-2004	(TREMBLrel. 28, Last annotation update)
DE	Carboxylesterase-like urinary excreted protein.	
GN	Name=Caxxin;	
OS	Mus musculus (Mouse) .	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Miyazaki M., Yamashita T., Taira H., Suzuki A.;	

```

RESULT 6
Q81034
ID Q81034 PRELIMINARY; PRT; 545 AA.
AC Q81034;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DD 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DD Carboxylesterase-like urinary excreted protein.
PDE Name=Cauxin;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC C Mammalia; Theria; Carnivora; Fissipedia; Felidae; Felis.
NCBI_TaxID=9685;
[1]
SEQUENCE FROM N.A.
FRN
RP

```

RC TISSUE=Kidney; PubMed=12401131; DOI=10.1042/BJ20021446;
 RA MEDLINE=22458314; Kamie K., Soeta S., Taira H., Yamashita T.;
 RT "Molecular cloning and characterization of a novel carboxylesterase-
 RT like protein that is physiologically present at high concentrations in
 RT the urine of domestic cats (Felis catus).";
 RL Biochem. J. 370:101-110(2003).
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB045377; BAC22577.1; -.
 DR HSSP; P12337; 1K4Y.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR000379; Ser esterase.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 DR KW Hydrolyase.
 SQ SEQUENCE 545 AA; 60505 MW; 9F73FA693D271FA9 CRC64;
 Query Match 64.5%; Score 1985; DB 2; Length 545;
 Best Local Similarity 71.7%; Pred. No. 6,5e-143;
 Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;
 QY 31 GPSAEGPQNRTRLGWIGQKQVTVLGSPPVNVVFLGVPAPPLGSLRFTNPQPASPMWNL 90
 DB 25 GPADAPVRSYTRLGKVRGKQTTVLGASTVPMVNLGIPYAAPPLGLRFPKQPPALPGNDF 84
 QY 91 REARSPYRLCQNSEFWLLDDHMLKVRPKFQVSDCLINITYAPAHADTSKLPVLYWF 150
 DB 85 RNARSYPLCLCQDLEWLVSYQHVLRPKLEASDCLYINITYAPAHADNCSNLPVWVWF 144
 QY 151 PGCAFPGKSGASIFPGSALAAVEDVLYVVVQYRLGIFGFTTMDQAPGNMAFKQVALLS 210
 DB 145 PGCAFPGKSGASIFPGSALAAVEDVLYVVQYRLGIFGFTTMDQAPGNMAFKQVALLS 204
 QY 211 WVQKNIEFFGDDPSVITIFGSSAGASISVSLISPMWAKLFRKAIMESGVAIIPYL-BAH 269
 DB 205 WVRNINIEFFGDDPSVITIFGSSAGASISVSLISPMWAKLFRKAIMESGVAIIPYL-BAH 264
 QY 270 DYKSEBDLQVVAHFCGNASDSEALRLRTRKPSKELLTSQTKTSFRRVVDGAFPPNEP 329
 DB 265 GDERKQDLQVVARJCGCHASDSALLQCLRAKPSBELMDISKULTFSIPVDDFFPEP 324
 QY 330 LDLSOKAFKAIPTIIGVNHCEGFLPMKEAPETLSGNSKSLALHLQNLHPIPOYLH 389
 DB 325 VALTQKAFNSVPSIIIGVNHCEGFL-STEFSBILGSSNSKSLALYVHTLNLNPTOYLH 383
 QY 390 LVANEYFPHDKSLTEIRDSLDDLGDVFFVVPALITARHRDAGAPVYFERRRPPQCFE 449
 DB 384 LVADHYFYNKSPVEIRDSFLDLDGDLVFPVGVVTAHYHDDAGAPVYFERRRPPQCLN 443
 QY 450 DTQKAPVKAHADDEVRYFGGAFPLKGDIVMEFEGATEEKKLSRKKMKYATPARTGNDG 509
 DB 444 DTRAPAFVADHSDERIRFVGGAFLKGDIVMEFEGATEEKKLSRKKMKYATPARTGNDG 503
 QY 510 NDLSLMPAYNLTBOYLQDLNMSLQGRKCEPRVDPWTSTI 549
 DB 504 EGVPLMPAYTQSEQYLKDLDSVSGQKKEQEVERPMWNTI 543
 RESULT 7
 Q95KH3 PRELIMINARY; PRT; 361 AA.
 AC Q95KH3;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Symonolus monkey).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecinae; Macaca.
 NCBI_TaxID=9541;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Temporal lobe right;
 RA Oseada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB060873; BAB46884.1; -.
 DR HSSP; P12337; 1K4Y.
 DR InterPro; IPR002018; Carboxylesterase.
 DR Pfam; PF00135; Coesterase; 1.
 DR KW Hypothetical protein.
 SQ SEQUENCE 361 AA; 41032 MW; AC62AB6E78C5F1A1 CRC64;
 Query Match 54.4%; Score 1676.5; DB 2; Length 361;
 Best Local Similarity 95.0%; Pred. No. 1.4e-119;
 Matches 323; Conservative 5; Mismatches 11; Indels 1; Gaps 1;
 QY 242 ILSPWAKGLFHKAIMESGVAIIPYLEADHYKESDQLQVVAHFCGNASDSEALRLRCLRTK 301
 DB 17 ILSPWAKGLFHKAIMESGVAIIPYLKAYDYKESDQLQVVAHFCGNASDSEALRLRCLRTK 76
 QY 302 PSKELLTSQTKTSFRRVVDGAFPPNEPDLISQKAFKAIPTIIGVNHCEGFLPMKEA 361
 DB 77 SSKEILLTSQKAKSFTRVVDGAFPPNEPDLISQKAFKAIPTIIGVNHCEGFLPM-BA 135
 QY 362 PEILSGNSKSLALHLQNLHPIPOYLHVNEXYFHDGSLTEIRDSLDDLGDVFFVVP 421
 DB 136 PEILSGNSKSLALHLQNLHPIPOYLHVNEXYFHDGSLTEIRDSLDDLGDVFFVVP 195
 QY 422 ALITARYHRDAGAPVYFERRRPPQCFEDTKAPAVKADHADVRFVFGAFLKGDIVME 481
 DB 196 ALITARYHRDAGAPVYFERRRPPQCFEDTKAPAVKADHADVRFVFGAFLKGDIVME 255
 QY 482 GATEEKKLSRKKMKYATPARTGPNNGNDLSLMPAYNLTBOYLQDLNMSLQGRKLEPR 541
 DB 256 GATEEKKLSRKKMKYATPARTGPNNGNDLSLMPAYNLTBOYLQDLNMSLQGRKLEPR 315
 QY 542 VDFWTSITPLILSASDMLHSPSLSLTLTLQPPFFCAP 581
 DB 316 VDFWTSITPLILSASDMLHSPSLSLTLTLQPPFFCAP 355
 RESULT 8
 SASB_ANAPL STANDARD; PRT; 557 AA.
 ID SASB_ANAPL;
 AC Q04791;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Fatty acyl-CoA hydrolase precursor, medium chain (EC 3.1.2.-)
 DB (Thioesterase B).
 OS Anas platyrhynchos (Domestic duck).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
 NCBI_TaxID=8839;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-65.
 RC TISSUE=Urological gland;
 RA MEDLINE=93300823; PubMed=8314791;
 RX Hwang C.-S., Kolatukudy P.B.;
 RT "Molecular cloning and sequencing of the thioesterase B cDNA and
 RT stimulation of expression of the thioesterase B gene associated with
 RT hormonal induction of peroxisome proliferation.";
 RL J. Biol. Chem. 268:14278-14284(1993).
 CC -1- FUNCTION: Fatty acid biosynthesis chain termination and release of
 CC the free fatty acid product is achieved by hydrolysis of the thio
 CC ester by a thioesterase. This thioesterase may be associated with
 CC peroxisome proliferation and may play a role in the production of
 CC 3-hydroxy fatty acid diester pheromones.
 CC -1- TISSUE SPECIFICITY: Highest levels in urological gland, much lower
 CC in liver and kidney.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L05493; AAA49223.1; -
 DR PIR: A47162; A47162.
 DR HSSP: O77540; 1K4Y.
 DR InterPro: IPR002018; Carboxylesteraseb.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 KM Direct protein sequencing; Fatty acid biosynthesis; Hydrolase; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 557 Fatty acyl-CoA hydrolase.
 FT ACT_SITE 227 227 Acyl-ester intermediate (By similarity).
 FT ACT_SITE 345 345 Charge relay system (By similarity).
 FT ACT_SITE 460 460 Charge relay system (By similarity).
 FT DISTLFD 93 122 By similarity.
 FT CAROHD 476 476 N-linked (GlcNAc...) (potential).
 SQ SEQUENCE 557 AA; 61637 MW; 03E3509A037E6B0 CRC64;

Query Match 42.2%; Score 1298.5; DB 1; Length 557;
 Best Local Similarity 48.4%; Pred. No. 1.9e-90;
 Matches 265; Conservative 79; Mismatches 177; Indels 27; Gaps 8;

QY 30 TGPSEAGQKTRKLCWQKQVTVGSPVNVFLGVPAPAPGLSLFTNPQAPSPWDN 89
 Db 23 TGQKQGEVVTNYSVGVYQVKNVAERSVVFGLFPAPKPVPLTFSPQPEPWKG 82
 QY 90 LREATSYNLCIONSEWMLDQHM-----KVNHYKFGVSDCYLNIYAAHADTQSKL 144
 Db 83 VRDAASYPMCLQDK---VLQGYLSDATNTRKVKRLQISDCYLYNTTSTVEDEKL 139
 QY 145 PVLVWFGAGFKTGSASIFDGSALAAVEDVLVWVQYRLGIFGFTTMDQAPGNMAFKD 204
 Db 140 PVFWVHGGGLVSGAASSYDGSALAFDNNVVVTIQYRLAGVFGSTGDKARNGMGLD 199
 QY 205 QVALLSWQKRIEFGGDSSTVTFGBSAGALSVSSLSLSMAGLFRKAIWEGVAL-I 263
 Db 200 QVALLOMIQENTIHFGDPSVTLFGBSAGVSVSALVSLPLAKGLFRKAISSGTVARI 259
 QY 264 PYLDAHDKESDQVVAHFCGNNASDEALRLCLRTKPSKEL--LTLQKTSFTVVD 321
 Db 260 LFTL-----QPEEQQRILAAAGCKSSSALVVECLREKTEHMEQITLKNPMTISLSD 315
 QY 322 GAFPPNPDLDSOKAKAITSITGVNNHCEGFLP-MKEAPEILSGSNKSLAHLIQT 380
 Db 316 GVFFPKSPRLSLSKVINAVYIIIGVNNCEFGMILPMMKFPEFTLEGKDVARQVLTST 375
 QY 381 LHI-----PQYLHLVANEYFHDGSLTEIRDSLTDLLGVFFVPPVPLIARVHDA 435
 Db 376 LALSFKAPEDIVLVNVEYIGVANEKRAQVADGLDIDPLPFLFSAVEVAHRHRRGNP 435
 QY 436 VYFYEPPHPOCFEDITKPAFVKADHAEVRFVFGAGLAKDIDVWEGATEEKLRSKM 495
 Db 436 VYFYEPPHPSAAGVVPFVKADHAEIAFVFGKPLAQN-----ATEEAKLSTTW 489
 QY 496 KYWATFRTGNPNPNDLSWPAKYLTEQYQLDLNLSLGRLEKPRVDFTSTITPLISA 555
 Db 490 KYTNFARNENPGEGLVHPQYDMERYLEIDLTKQAKKLEKRMKEFMQVLTQEI MSD 549
 QY 556 SDMLASPL 563
 Db 550 RRRKHTDL 557

RESULT 9

091WGO
 ID 091WGO PRELIMINARY; PRT; 561 AA.
 AC 091WGO;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 26, Last annotation update)
 DE Carboxylesterase 2 (Carboxylesterase M3 precursor).
 GN Name=Ces2; Synonyms=ces2a3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon, and Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Yoshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smalios D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RL Strausberg R.;
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RL Strausberg R.;
 RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RL Strausberg R.;
 RN Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon, and Kidney;
 RL Strausberg R.;
 RN Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=22744052; PubMed=12859986; DOI=10.1016/S0003-9861(03)00286-8;
 RA Furukawa T., Hosokawa M., Nakata F., Sacoh T., Chiba K.;
 RT "Purification, molecular cloning, and functional expression of
 RT inducible mouse liver acylcaritine hydrolase in C57BL/6 mouse,
 RT belonging to the carboxylesterase multigene family.";
 RL Arch. Biochem. Biophys. 415:101-109 (2003).
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL: BC015290; AAH15290.1; -
 DR EMBL: BC024552; AAH24552.1; -
 DR EMBL: BC031170; AAH31170.1; -
 DR EMBL: BC034178; AAH34178.1; -
 DR EMBL: BC034180; AAH34180.1; -
 DR EMBL: BC034191; AAH34191.1; -
 DR EMBL: AB110073; BAC76623.1; -

DR HSSP; P12337; 1k4Y.
 DR MGD; MGI:2385905; Cess2.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 DR HydroLase; Signal.
 FT SIGNAL 1 26 Potential.
 FT CHAIN 27 561 carboxylesterase ML3.
 SQ SEQUENCE 561 AA; 62469 MW; P793967A646EB728 CnC64;
 Query Match 41.7%; Score 1282.5; DB 2; Length 561;
 Best Local Similarity 48.4%; Pred. No. 3.2e-89;
 Matches 276; Conservative 77; Mismatches 174; Indels 43; Gaps 14;
 QY 14 FELLIQPLGRGKGTGPSAGPQNRRLGMIQKQVTVGSPVAVVVFPPRAAPL 73
 DB 15 FGLLL--LLHVQ--GQDSPEA--NPIRNTHTQIQSLIHVADTKA GHTPLGIPAKDPV 70
 QY 74 GSLRFTNPQAPSPMDNLEATSYPLCLQNSEWL----LLDQHLKVHYPKGVSEDCLY 129
 DB 71 GPLAFAPPEAPBPPMGSVADGTAAHPMCLQNDMLNEAGLPMMKMLSSFP---NSEDCLY 127
 QY 130 LNIYAPAHADTGSKLPLVLPFGGAFKTSASIFDGSAIAAYEDVLVVVVGRIIGIFGF 189
 DB 128 LNIYTPAHAGHGSNLPVMMWHIGALVIGMASMEFGSLITVNEIDLVTITIGRLGVLGFF 187
 QY 190 TTMQOHAPGNNAFPDOVAALSNVQKNIFRFGDPSVTITRFSAGALIVSSLLISPMKAG 249
 DB 188 STGQOHARGNMGYLDQAAALRMVQONLHFGGNPRVTIFGSSAGTSSSHVSPMSQG 247
 QY 250 LFKHAINESGVAIIPYEADHYEKSEDLQ--VVAHCGNNADEALRLCLRTKSEKELLT 308
 DB 248 LFKHAINESGVALPDLIS---ETSENVSTTVAKLSGEANDSQALVCLGKSGAEITLA 304
 QY 309 LLSQTKSFTRVVDGAFPPNEPDLISQAKFAIPSIIGVNNHCEGFLPMKEAPEILSGS 368
 DB 305 INKVFKNIPAVVDEFPFRHPEKELLASDFHPVPSIIIGVNDDEFQMSIP-----VVMGS 358
 QY 369 N---KSLALHLIQNHL-----IPROYLHLVANEVHYDGHSLTEIRDSLLDLGDVFF 418
 DB 359 AQMIKGITRENLQVLDKDTAVQMLPPECSDLMEETGDEDAQTQIDQFTENMGDFWF 418
 QY 419 VVPALITARYRDGAPYFYEFRRRPOCFEDTKPAFYKADHAEVRVFGGAF--LKGD 476
 DB 419 VIPALQVNAHQR--SHAPYFYEFQHPBSYFKDVRPRHYKADHAEIRPVVVSFFPMKMLD 477
 QY 477 IVMEGATEEBEKLRSKKMKYATFARTGNGNDLSLMPAYNLTEOYLQDLNMSLQQR 536
 DB 478 F-----TEEBELLSRRMKYMANPARHGNNSGLPYVPMWDHDEOYLQDIPAVGRA 531
 QY 537 LKEPRVDFWTSTIP---LILSASDMLSP 563
 DB 532 LKAGRLQFWTKTLPOKIQELKASQDKHREL 561
 RESULT 10
 Q8BK48 PRELIMINARY; PRT; 559 AA.
 AC Q8BK48;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
 DE RIKEN full-length enriched library, clone:503151919 product:similar
 DE to CARBOXYLESTERASE (EC 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE)
 DE (MONOXYTRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
 DE (9030624102Rik protein).
 GN Name=9030624102Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Azawa K., Nagao S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama T., Nishi K., Katsunuma T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasahagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsushita S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX Adachi J., Azawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka I.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Klausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Umeda T.B., Toshikiyaki S., Carninci P., Prange C.,
 RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP [6]
RN SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AK077248; BAC36707.1; -.
DR EMBL; BC055062; AAH55062.1; -.
DR HSSP; P12337; 1K4Y.
DR MGI; 2443170; 9030624L02Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR002018; Carboxylesteraseb.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolyase.
SQ SEQUENCE 559 AA; 62317 MW; CC14C61034A122C3 CRC64;

Query Match 41.6%; Score 1281; DB 2; Length 559;
Best Local Similarity 48.7%; Pred. No. 4.2e-89;
Matches 260; Conservative 79; Mismatches 165; Indels 30; Gaps 8;

QY 31 GPSAGPQKNTLGIQKQVTVLSPVNVNVLGVPAAPLGLRTNQPASPMNL 90
DB 26 GQDASPIRNTHTGQVRSLVVKTDLAVHTFLGIPFAKPPVGLRFAPEPAPMGV 85
QY 91 REATSYNLCIONSEWL-LDQMLKVHPKFGVSEDCLYINIYAPAHADTGSKLPLV 149
DB 86 RDGTSHPMNCIONDMIMGSEDLKMMNLILPISMSBDCLYINIYAPAHAGSNLPLV 145
QY 150 FPGGAFKTSASIFPGSALAAVEDLVVNVQYRLGIFGFTTMDQAHAGNAFKDQV 209
DB 146 IHGALTYGASMSDGMALATEDVVVAIQYRLGLVGFSTGDOAHAKMGVYLDQV 205
QY 210 SWVQKNIEFGSDPSVTIFGSSAGALISVSLISPMKGLFHAKIMSGVAILPYL 269
DB 206 RMVQONIVHFGSDPRTVITFGESAGTSSVSHVSPMSGGLFHAKIMSGVAILP 265
QY 270 DYKSESDQVVAHFCGNNADEALLRLCTRTKPSKELLTTSQKTSFTRVVDGAF 329
DB 266 SSEVNH--RIYANISGCAAVNSETLMLCKRKNEMLAINKVKIIPGVVDGEFLPK 323
QY 330 LDLSQKAFKAPISIIIGNNHECGFLP-----MKEAPEIISGNSKSLALH 376
DB 324 QELMASDQFHVPISIIIGNNDEYWIIFPTMDPQKTEITRKLPVLV-----K 379
QY 377 IONLIHLPQYLVHANEYFHDKSLTEIRDSLLDLGVFVVPALITARYRDGAPV 436
DB 380 M-----LPFGSGDLIMEYMGDTEDPELLOAFEMKGMDFMFIPLNLOVANH 433
QY 437 YFYEFRRHPCCFEDTKAPFKADHAEVRFVFGAFLKGDIVMFEAGTEEEKLSR 496
DB 434 YFYEFRRHPCCFEDTKAPFKADHAEVRFVFGAFLKGDIVMFEAGTEEEKLSR 489
QY 497 YMATFARTGNPNNDLSLWPAVNLTEOYLOLDLMSIGRLKEPRVDFTWSTIP 550
DB 490 YMATFARTGNPNNDLSLWPAVNLTEOYLOLDLMSIGRLKEPRVDFTWSTIP 543

RESULT 11
035533

ID 035533 PRELIMINARY; PRT; 559 AA.
AC 035533;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Carboxylesterase precursor (EC 3.1.1.1).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian golden; TISSUE=liver;
RA Some T., Wang C.Y.;
RL "Microsomal amidases and carboxylesterases.";
RL (In) Guengerich F.P. (eds.),
COMPREHENSIVE TOXICOLOGY VOLUME 3, Biotransformation, pp.265-281,
Pergamon, Oxford (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian golden; TISSUE=liver;
RA Some T., Ishida Y., Takabatake E., Wang C., Isobe M.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; D50577; BAA23605.1; -.
DR HSSP; P12337; 1K4Y.
DR GO; GO:0004091; F:carboxylesterase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR InterPro; IPR002018; Carboxylesteraseb.
DR InterPro; IPR000866; ER_target_S.
DR InterPro; IPR000379; Ser_estrs.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Hydrolyase; signal.
FT SIGNAL 1 27 potential.
FT CHAIN 28 559 carboxylesterase.
SQ SEQUENCE 559 AA; 62126 MW; 5B6B4F9A47C6AD8 CRC64;

Query Match 41.5%; Score 1276.5; DB 2; Length 559;
Best Local Similarity 49.6%; Pred. No. 9.2e-89;
Matches 262; Conservative 79; Mismatches 172; Indels 15; Gaps 9;

QY 28 GKTGPABEPQKNTLGIQKQVTVLSPVNVNVLGVPAAPLGLRTNQPASPM 87
DB 26 GQDPS--PIRNTHTGQVRSLVVKTDLAVHTFLGIPFAKPPVGLRFAPEPAP 82
QY 88 DNLREATSYNLCIONSEWL-LDQMLKVHPKFGVSEDCLYINIYAPAHADTGS 146
DB 83 SGVRDGNSEPMNCIONDMIMGSEDLKMMNLILPISMSBDCLYINIYAPAHAG 142
QY 147 LVMPFGGAFKTSASIFPGSALAAVEDLVVNVQYRLGIFGFTTMDQAHAGNA 206
DB 143 MWYINGALVVGMAVMYDGSMLATEDVVVAIQYRLGLVGFSTGDBHAGNGVY 202
QY 207 AALSWQKNIEFGSDPSVTIFGSSAGALISVSLISPMKGLFHAKIMSGVAILPY 266
DB 203 AALRWQONIVHFGSDPRTVITFGESAGTSSVSHVSPMSGGLFHAKIMSGV 262
QY 267 EAHDYKESDQVVAHFCGNNADEALLRLCTRTKPSKELLTTSQKTSFTRV 326
DB 263 ISSSEEMV--YTIYANISDCAAVNTEVLVSLRKSGSEAILANKVKIIPAV 320
QY 327 NEPLDLSQKAFKAPISIIIGNNHECGFLP-----MKEAPEIISGNSKSL 382
DB 321 KHPQELMASDQFHVPISIIIGNNDEYWIIFPTMDPQKTEITRKLPVLV----- 380
QY 383 IIPQYLVHANEYFHDKSLTEIRDSLLDLGVFVVPALITARYRDGAPVYFEFR 442
DB 381 LPFGSGDLIMEYMGDNEDPELLOAFEMKGMDFMFIPLNLOVANHFOR--AA 439
QY 443 HRPCCFEDTKAPFKADHAEVRFVFGAFLKGDIVMFEAGTEEEKLSRMMKY 502


```

Db 440 HRSPKOTKTPHYADHGDSEFFFGN-LLOG--IKR-ASTEEBELLSRKKMKWAFNA 495
Qy 503 RTGNPNNDLSLMPAYNLTEQYLQDLNMSLGRKKEPVDFTSTIP 550
Db 496 RHGNPNSENLFPYWPWLDHDEQYLDIDIKPAVGKALKARRLOFTWTLIP 543

RESULT 12
Q8BM97 PRELIMINARY; PRT; 559 AA.
ID Q8BM97;
AC Q8BM97;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Mus musculus adult male colon CDNA, RIKEN full-length enriched
DE library, clone:9030624102 product:similar to CARBOXYESTERASE (EC
DE 3.1.1.1) (ALI-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
DE (PROCAINE ESTERASE) (METHYLBUTYRASE).
GN Name=9030624102Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=99279253; Pubmed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=21085660; Pubmed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20499374; Pubmed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20530913; Pubmed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama U., Nishi K., Katsuna T., Teshiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Kashiwagi K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

```

```

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayano K., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori P., Imotani K., Ishii Y., Itoh M., Kagawa I., Katsukawa T.,
RA Katoh H., Kawai U., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsumoto T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akibira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AK033563; BAC28361.1; -.
DR HSSP; P12337; 1k4Y.
DR MGD; MGI:2443170; 9030624102Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR002018; Carbestereab.
DR InterPro; IPR000179; Ser_estra.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYESTERASE B_1; 1.
DR PROSITE; PS00941; CARBOXYESTERASE_B_2; 1.
KM Hydrolyase.
SQ SEQUENCE 559 AA; 62305 MW; 8C54D21522C352C4 CRC64;

Query Match 41.3%; Score 1271; DB 2; Length 559;
Best Local Similarity 48.7%; Pred. No.2.4e-88;
Matches 260; Conservative 78; Mismatches 166; Indels 30; Gaps 8;

Qy 31 GPSAEGPQNRTRLGWIGQKQYTVLGSPVPVNVFGLVPPAAPPLGSLRFTNPQAPSWNL 90
Db 26 QGDSASPIRNTHTGQVRSGLVHVVKDIDYAVHTFGLIPAKKPVGRLPAPAPAPMAGV 85
Qy 91 REATSYPLCLQNSEWL-LDQHLKHYKPKFGVSEDCLYNINYPALADGSLPVLVW 149
Db 86 RDGTSHPMWCQONNLMGSEDLKMMNLLPRTMSSEDCLYNINYPALAHGSLNPVWVW 145
Qy 150 FPGGAFKTSASIFPGSALAYEDVLVWVQYRGIFPGFTTPOHAGNNAFKQVVAL 209
Db 146 IHGHALTVGASMTDGSMLAATEDVYVAIQRLGVGFSTGDAHAGNMGYLDQVALL 205
Qy 210 SWVQKNIEFFGSDPSVTYIFGESAGAISVSSLISPMKGLFHKAIMESGVAIIPYLEAH 269
Db 206 RMVQONIVHFGSNDPDRVTYIFGESAGTSVSHVVSPMGGLFHGAIMESGVAIIPDLISS 265
Qy 270 DYKSESDIQVYAHCCGNNAASEALRLCKTRKPSKELLTSQKTSFRRYVDGAFFNBP 329
Db 266 SSEWVH--RIVANISGCAAVNSETLMCCLRGKNEAEMLAINKVFRTIPGVVDGEFLPHP 323
Qy 330 LDLSQKAFKAIPSTIIGVNNHECGFLP-----MKEAPEILSGSNKSLALHL 376
Db 324 QELMASKDFHFPVPSITIGINDBYGMILPTIMDPACKIEITRKILPAVL----KSTALKM 379
Qy 377 IONILHIPPQYLIHVAHYFPHDKSLTEIRDSLDDLGDVEFVVVPALITARYHDAGAPV 436
Db 380 M-----LPBEGDDLMEETMGDTEDPETLQAGPFEMKMDPMFVLPALGVAFQR-SHAPV 433
Qy 437 YFEYFRHRPQCFEDTKPAFVADHADAVRYVFGCAFLLKGLDIWMFGATBEKLSRKKM 496
Db 434 YFEYFQHRPSFFKXFRPPYVADHGDDEIFLVGFQF--GNIKL--PYTEEBEQLSRMMK 489
Qy 497 YMAFPAFTGNPNNDLSLMPAYNLTEQYLQDLNMSLGRKKEPVDFTSTIP 550
Db 490 YRANFAHNGNPNSEGLPYWPWLDHDEQYLDIDIPSGVGRALKARRLOFTWTLIP 543

RESULT 13
Q8PDB7 PRELIMINARY; PRT; 556 AA.
ID Q8PDB7;
AC Q8PDB7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein BC015286.

```

GN Name=BC015286;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL; BC058815; AAHS815.1; -
 DR HSSP; P12337; 1K4Y. -
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR003379; Ser. esters.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 DR HydroLase; Hypothetical protein.
 KW SEQUENCE 556 AA; 61927 MW; 58DE05A33C73BECD CRC64;
 Query Match 41.1%; Score 1266.5; DB 2; Length 556;
 Best Local Similarity 47.1%; Pred. No. 5.3e-88;
 Matches 273; Conservative 76; Mismatches 190; Indels 41; Gaps 12;
 QY 1 MPOGLTSSASQWCFPLI--LQPLIGHQMGKTSABEPQNTKLGMIQKQVTVLGSV 58
 DB 1 MPR--SOMNMWLVLLFGLLLGHVQ-GQDSPEA-SPIRNTTGGVRSGLVHVTKTKA 55
 QY 59 PNVNVLGVPAAPPLGSLFTNPQASPMWDLREATSYPNLCIONSEMLLDQMLLVHY 118
 DB 56 GVHFLGLGIPFAKPPVGLRFAPEAPBPMWGSVGRGTNAHMACLN---LGVKKEITGLK 111
 QY 119 PKRGVSDCLYNTIYAPAHADTGSKLPLYVWFGCAFKTSASITFDGSSALAAVEDLVVV 178
 DB 112 PNVSTSDCLYNTIYTAHHAHSGNLPEVWVWVHGGGLVAGMASWYDGLLAIEDLVVVV 171
 QY 179 VOYRLGIFGFFTTWDQAPAKFOVAALSWQKRIEFGDGPSVITFGSSAGASIV 238
 DB 172 IOYRLGVLFGFSTGQDQARNGMGLDVAALRWIQQNIHFGGPPDVITFGSSAGGTSV 231
 QY 239 SSLLISPMAGLPHKAIMESGVAIIPYLEADHYEKSEDLQVNAFCGNNADESEALLRCI 298
 DB 232 SSHVSVSMXKGLFRGAIMESGVALPLYLITDTSBMVS--TTVAKLSCGCAEMDEALVRCL 289
 QY 299 RTSPSKBLILTSQKTSFTTVNDGAPFPNPPLLSQKAKAIPSTIGVNNHCEGLFLPM 358
 DB 290 RGSSEBILILNKLVQMTIPAVNDGEFPFPHKELASDEHPVPSIIGVNNDEFGWTID- 348

QY 359 KEAPILSGNSKSLALHIONI-----LHIPPOLYLVANEXFHDKSLTEIRDS 408
 DB 349 ----VWVGSAGQTIKEIRRENLQVLKNTTQNLMPBCSLLMEFTMGDTEDQTIQ 403
 QY 409 LLDLLGDVFPVVPALITRHYRDGAPYFYEFPFHRQCFEDYKPAFVKADHAEVAFV 468
 DB 404 FTEWMEDEWFVIALQVAVYFQR-SHASYFYEFOHOIASLKDVAFTVKADHAEIPFV 462
 QY 469 GGAF--LKGDIWMEGATEEBEKLRSKMMKYMATFARTGNPNNGMDLSLWPAVYNTEQY 526
 DB 463 GYFFMDKMLD-----TEGKILSRMMKMYAFNAGHNPNSGGLPYWPMVDHDEOYLQ 516
 QY 527 LDLMNSLGQRLKEPRVDFWTSTIP--LILASDMLHSP 563
 DB 517 LDTPAVGRALKSRRLQFWFTKLSQKIQLERASQDKTEL 556
 RESULT 14
 ID OCGM54 PRELIMINARY; PRT; 568 AA.
 AC OCGM54;
 DT 05-JUL-2004 (TREMBlrel, 27, Created)
 DT 05-JUL-2004 (TREMBlrel, 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel, 27, Last annotation update)
 DE LOC443703 protein (Fragment).
 GN Name=LOC443703; (African clawed frog).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius U., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RC Klein S., Gernard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL; BC074230; AAH74230.1; -
 DR GO; GO:0016787; F:hydrolase activity; IEA.

DR InterPro: IPR002018; CarboxylesteraseB.
 DR InterPro: IPR000379; Ser esterase.
 DR Pfam: PF00135; Coesterase_1.
 DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 DR HydroLase.
 DR KMW
 FT NON TER
 SEQUENCE 568 AA; 63032 MW; 9336D08B53931A1 CRC64;

Query Match 41.0%; Score 1263.5; DB 2; Length 568;
 Best Local Similarity 48.2%; Pred. No. 9.2e-88;
 Matches 258; Conservative 87; Mismatches 169; Indels 21; Gaps 8;

QY 28 GKTGPSAAGPORNRLGWIQKQVTVLGSFVNVVFLGVPAPPLGSLRTNPOAPPM 87
 DB 28 GIYTEDAPRLITNYGQLGKTVGAKETDRLIHVFMKVPAPKAPDIPGLAREDEPQPEPM 87
 QY 88 DNLREATSPNLCIONSEWL--LLDQHLKVHYPKFVSEDCLYLNTIYAPAHADTGSKL 145
 DB 88 SSIREATENPMICQDKKMGQLADFFAKDPFP--VSEDCLYLNTVPTPADRGNEPLP 145
 QY 146 VLVPFGGAFRTGSASIFDGSALAYEDVLVVVVOYRLGIFGFTTWDQHPGKMAFKDQ 205
 DB 146 VNVFIHGGLTGMGAGMEGSGALAYENVVVSIOYRLGIMGFSTGDKKARGNYGFLDQ 205
 QY 206 VVALSWQKNIIEFGGDPSSVTIFGESAGATSVSSLLSPMAKGLFHKAIMESGVALIPY 265
 DB 206 VVALRWNRDNKIDFGNGQSVTIFGESAGLSVSAQVLSPLSKGLFHKAISSGVALIPG 265
 QY 266 LEADHYEKSEDLQVAVAFCCGNNAADS--BALRLCLRTKPSKELLTLGQTK--SFTPVVDG 322
 DB 266 LMA---SKTEKILRLIHVVAISSCSVSLADCLKKTEDEIIVAIISAMKVAFAVVDG 322
 QY 323 AAPPNEPLDLISQAFKAIPSIIGVNNHCCFLLPMKEAPILSGNSKSLAHLIQLNTH 382
 DB 323 VFLPKPAEELIASKESNPVPELIGVNNHFCMILPLALN--ISGYREGMKDIOSTLV 379
 QY 383 IIPQYLA-----LVANVEPHDKSLTEIRDSLIDLDGVPFVVPALITARYADAP 435
 DB 380 ALP-FVHSFTSVVPEIMEEYGDNDPKELNNFNLVDGDIIFVLPALRTKXHYHDSHP 438
 QY 436 VYFEPFHRPQCFEDTKPAFAFKADHAEVRFVFGGAFPKGDIIVMEGATEEBSKLSRRKM 495
 DB 439 VYFEPFHRPBMVDSKDDPKADHGBDELYFVVGPFKSGILFRKSNTEBSKLSKTIIM 498
 QY 496 KYMAFPAATGNGNDLSLMPAYNLTEQYLOLDLMSIGQLKEPRVDFTSTIP 550
 DB 499 KYMAFPAATGNGNDLSLMPAYNLTEQYLOLDLMSIGQLKEPRVDFTSTIP 553

RESULT 15
 070177 PRELIMINARY; PRT; 561 AA.

AC 070177;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Carboxylesterase precursor (EC 3.1.1.1).
 GN Name=carboxylesterase;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=small intestine;
 RA Sore T., Kunitomo T., Isobe M.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB010632; BAA55691.1; -.
 DR HSSP; P37967; IOE3.
 DR GO; GO:0004091; F:carboxylesterase activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR InterPro: IPR002018; CarboxylesteraseB.
 DR InterPro: IPR000379; Ser esterase.
 DR Pfam: PF00135; Coesterase_1.
 DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 DR HydroLase; Signal.
 FT SIGNAL
 FT CHAIN
 SEQUENCE 561 AA; 62239 MW; 73A468C33F969398 CRC64;

Query Match 41.0%; Score 1262.5; DB 2; Length 561;
 Best Local Similarity 48.8%; Pred. No. 1.1e-87;
 Matches 275; Conservative 77; Mismatches 179; Indels 33; Gaps 13;

QY 1 MPQG-LTSSASQMKFLLIPLGLHRQWKTPGSAEGQRTNRLGKQVTVLGSFV 59
 DB 1 MPRQLHSMVLAIVFGLL--LLIHVQ-GQDSPPS-SPIRTTHGQVQGLDHDVDTAG 56
 QY 60 VNVFGLVFPAPPLGSLRTNPOAPSPMDNLEATSPNLCIONSEWL--LLDQHLK 115
 DB 57 VHTLGLIFPAKPPGPIRFAPPEPEPMSGVADATQPMCLQDLIDEGGLDMKMI- 115
 QY 116 VHYPKFVSEDCLYLNTIYAPAHADTGSKLPLYVMPFGGAFRTGSASIFDGSALAYEDVL 175
 DB 116 --LSSISMSSEDCLYLNTIYAPAHAREGSNLPVWVIHGGLVVGMAWMDGSLITVNEDLV 173
 QY 176 VVVVOYRLGIFGFTTWDQHPGKMAFKDQVVALSWQKNIIEFGGDPSSVTIFGESAG 235
 DB 174 VVTIOYRLGVYGFSTGDEHARGMWGLDQVVALRWQNLIAHFGNGNRYITIFGESAG 233
 QY 236 ISVSSLLSPMAKGLFHKAIMESGVALIPYLEADHYEKSEDLQ--VVAIFCCGNNAADSSEAL 294
 DB 234 TVSSSHVISPMSQGLFHGAIMESGVALIPDLIS--ETSEVSTTVAKLSCSEAMDSBAL 290
 QY 295 LRCLRTKPSKELLTLGQTKSFTRVVDGAFPEPNELDLGKAFAKAIPSIIGVNNHCCGF 354
 DB 291 VRCLRAKSGAELIVNVKFKMIPAVVDGEFLRHPKELLASEDPHPVPSIIIGNVTDEYCC 350
 QY 355 LIPM-----KAPILSGNSKSLAHLIQLNTHLIPQYLAHVANVEPHDKSLTEIR 406
 DB 351 TIIPWMTAOLIKK-----LSRENLAVALKDTAAQMLPPECGDILMEBYMONTDSDOTLQ 406
 QY 407 DSLIDLGDVFPVVPALITARYHDAGAPVYFESFRRPQCFEDTKPAFAFKADHAEVRF 466
 DB 407 IQYTEMGDFLFLVLPALQVAHFOR-SHAPVYFYFGAIPSTFKXVRPHVADADAEVRF 465
 QY 467 VFGGAFPKGDIIVMEGATEEBSKLSRRKMXYMATFARTGNGNDLSLMPAYNLTEQYLO 526
 DB 466 VF-GSPFSGMKLDE--TEBERLLSRMMKMYMAFAQGNBSGLPYWPAALDHDEQYLO 521

QY 527 LDLMMSIGQLKEPRVDFTSTIP 550
 DB 522 LDTHPAVDRAKARRLQFWTKTLP 545

Search completed: June 15, 2005, 09:39:21
 Job time : 95 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 09:25:21 ; Search time 25 Seconds
(without alignments)
2236.078 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079

Sequence: 1 MPGLTSSASQWCFLLQF.....PLSSLTPLSLQPPFFCAP 581

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298.5	42.2	557	2	A47162 thiolesterase B (E
2	1231.5	40.0	532	2	A34329 60K esterase (E
3	1230	39.9	559	1	JC5408 carboxylesterase (
4	1199	38.9	561	2	S47655 carboxylesterase (
5	1186.5	38.5	554	1	S34607 carboxylesterase (
6	1163	37.8	554	2	A39060 carboxylesterase (
7	1162.5	37.8	566	2	S19307 carboxylesterase (
8	1144	37.2	549	2	JX0054 carboxylesterase (
9	1140	37.0	567	1	A41010 carboxylesterase (
10	1139	37.0	561	2	S62788 carboxylesterase (
11	1138.5	37.0	562	2	A55281 carboxylesterase (
12	1136	36.9	540	2	A31584 carboxylesterase (
13	1129.5	36.7	561	2	S71597 carboxylesterase (
14	1123	36.5	561	2	JC2447 carboxylesterase (
15	1121.5	36.4	565	2	S10367 carboxylesterase (
16	1034.5	33.6	539	2	A29923 carboxylesterase (
17	777	25.2	596	1	ACRYE acetylcholinestera
18	764.5	24.8	614	2	A39256 acetylcholinestera
19	763	24.8	614	2	JH0314 acetylcholinestera
20	756	24.6	599	1	A38868 acetylcholinestera
21	754	24.5	614	2	JH0811 acetylcholinestera
22	752	24.4	602	1	ACHU acetylcholinestera
23	740	24.0	603	2	S70849 cholinesterase (E
24	732	23.8	581	2	C39768 cholinesterase (E
25	726.5	23.6	584	2	S48724 acetylcholinestera
26	710.5	23.1	583	2	S10712 acetylcholinestera
27	692.5	22.5	620	2	A54413 triacylglycerol 11
28	660	21.4	745	2	S13586 para-nitrobenzyl e
29	658.5	21.4	489	2	B69680

30	643.5	20.9	597	2	A33668 sterol esterase (E
31	633.5	20.6	599	2	A57701 sterol esterase (E
32	627.5	20.4	664	2	JC7990 acetylcholinestera
33	618	20.1	612	2	A34967 sterol esterase (E
34	594.5	19.3	691	2	JB0150 acetylcholinestera
35	593	19.3	550	1	A34576 crystal protein pr
36	586	19.0	767	2	S47639 acetylcholinestera
37	581.5	18.9	746	2	A25363 acetylcholinestera
38	575	18.6	540	2	S51043 carboxylesterase (
39	572.5	18.6	602	2	T37254 acetylcholinestera
40	570.5	18.5	637	2	S66236 acetylcholinestera
41	569.5	18.5	554	2	T31783 hypotetical prote
42	569.5	18.5	557	2	A56690 esterase - Caenorh
43	562.5	18.3	545	2	A89046 protein B0238.1 [i
44	557.5	18.1	593	1	S25062 triacylglycerol 11
45	551	17.9	562	2	S27782 esterase precursor

ALIGNMENTS

RESULT 1

A47162 thiolesterase B (EC 3.-.-.-) precursor - mallard
C.Species: Anas platyrhynchos (mallard)
C.Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004
C.Accession: A47162
R.Hwang, C.S.; Kolatukudy, P.E.
J. Biol. Chem. 268, 14278-14284, 1993
A.Title: Molecular cloning and sequencing of thiolesterase B cDNA and stimulation of expre
A.Reference number: A47162; MUID:93300823; PMID:8314791
A.Accession: A47162
A>Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-557 <HMA>
A.Cross-references: UNIPROT:004791; GB:J05493; NID:9213100; PIDN:AAA49223.1; PID:9213101
C.Superfamily: cholinesterase; cholinesterase homology
C.Keywords: hydrolase
P:56-545/Domain: cholinesterase homology <CHE>

Query Match 42.2%; Score 1298.5; DB 2; Length 557;
Best Local Similarity 48.4%; Pred. No. 4.1e-94;
Matches 265; Conservative 79; Mismatches 177; Indels 27; Gaps 8;

QY	30	TGSAEGQQRNTRGMIQKQVTVGSPVNVFVGPPAPRGLRFTNPQASPMDN	89
DB	23	TGQKAEQPEVNTGVSVRGQVKNAAERSVNVGLPFAKPVGLPSEQPEPMKG	82
QY	90	LRATSYENLCLQNSEWMLDQHML----KVHYPKFVSEDCLYLNTYAPAHADTGSXL	144
DB	83	VRDASVPMCLQDK---VLGQYLSDAITNRKKEKRLQISDCLYLNTYTPVSTEQEKL	139
QY	145	PVLVWFPQGAFTKTSASIFDGSALAAVEDVLVVVQYRLGIFGFTTWDQAHAGNMAEKD	204
DB	140	PFVWVIHGGGLVSGAASGYDSALAAPNVVVTIQLYRLAGYSTGDKHARGMWGYLD	199
QY	205	QVALASWQKNIIEFGGPPSSVTTTIGESAGALSVSLLSPAKLFLPKHAKMESGVAL	263
DB	200	QVALQWIOENIIRHFGPGSVTIFGESAGGSVALSLPLAKGLFKHAKSESSTAVRI	259
QY	264	PYLEAHVYKESBDLVVAHFGGNNASDSEALLRCRTKPSKEL--LTLSQKTSFTTRVVD	321
DB	260	LFTL---QPEEQARIRAAAGCESSAALVECLREKTBEMEGITLKMPPMFLSASLD	315
QY	332	GAFPNPEBLDLISQAFKAPISIIIGVNNHBCGFLP--MKEAPELISGNSKSLAHLQNI	380
DB	316	GVFPFKSPRQLLSEKVINAVPIIIGVNNCEGMIIPRMKKPEFTGEGKQVAAQVQLST	375
QY	381	LHI-----PQYLIHVNNEYTHDKSLTEIRDSLIDLDGVPFVVVPALITARYHDAQAP	435
DB	376	LALSFKGAPSDIIVDLVYNEYIGVANNRAQVBDGLDLSIADPLFVFSAAVEVAHHHDAQNP	435
QY	436	VYFVYFRHRRPQCFEDTKDAFVKADHAEVRFVFGGAFKGDIVMPEGATEEBKLSRRKM	495

```

Db      436 VYFVEFGHRSASAGVPEFVKADHDEIAFVGKPLAEN-----ATBEKSLSTTW 489
Qy      496 KYATATPRTGPNNGNDLSLMPAYNLTEQYIQLDINSLGRLKEPRVDWTSTIPILSA 555
Db      490 KYTNAPARNNGNPGEGVLVWPQYDMERYLEIDLQYAAKCLKERKMEFWMQLTEQIMSD 549
Qy      556 SDMLHSPL 563
Db      550 RRKHTDL 557

RESULT 2
A34329
60k esterase (EC 3.1.1.-) isoform 2 - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004
C/Accession: A34329
R/02018, U.
J. Biol. Chem. 264, 12533-12545, 1989
A/Title: Isolation, properties, and the complete amino acid sequence of a second form of
A/Reference number: A34329, MUID:89308686, PMID:2745458
A/Accession: A34329
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-532 <O20>
A/Cross-references: UNIPROT:P14943
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase
F.123-517/Domain: cholinesterase homology <CH>
F.201.430/Active site: Ser, His #status predicted

Query Match      40.0%; Score 1231.5; DB 2; Length 532;
Best Local Similarity 47.8%; Pred. No. 7-2e-89;
Matches 253; Conservative 76; Mismatches 167; Indels 33; Gaps 7;

```

```

Qy      37 PORNRLMIQKQYTVLGSVPVNVFLGPPAPPLGSLRFTNDOPASPMNDIREATSY 96
Db      6 PLRNHTQVGRSLVHVGSTIDAGVHTFLGIPAKPLPLRFAPEPAEASGVADGSL 65
Qy      97 PNLCLQNSEMLLLDQMLKMHY--PKFGVSEDCVLTNLYAPAHDTGSKLPVLVWPFQGA 154
Db      66 PAMCLQN--LAIMDDVLLHFTPTSPWSEDCVLTNLYSPAHAEESDLPVWVWHIGGG 123
Qy      155 FKTSASIFDGSALAAVEDLVVVVQYRLGIFGFTTWDQAPGNMAFKDQVAALSWYQ 214
Db      124 LTMGASWYDGSALAAFEDEVVVVTIQYRLGIFGFTTWDQAHGHTDQVALLRWYQ 183
Qy      215 NIEFGDPSVITFGESAGALSVSSLSLSPMAKGLFKHAIMESGVAILPYLEADHYKS 274
Db      184 NIAHFGNPGRVITLFGESAGTSVSHVLSPMISGLFGALMESLIVALLPGLITSSEVV 243
Qy      275 EDLVVAFHFCGNMADSSEALRLCLRTKPSKELTLISQTKSFTRVVGAFEPNPELILS 334
Db      244 S--TVVNLNLSRCGVSTTLVRCRAKSEEMLAITQVPMILPGVVGDFLPRHPEBLA 301
Qy      335 QAKAKALPSITGVNNHCEGFLP-----MKEAPEILSGNSKLALHLIONIL 381
Db      302 LADQAPVPSIIGINNDEGWTIPKLLAIDPQEBRDQAMEIIMHQTQKMLT----- 354
Qy      382 HIPOYLHLVANEYFHDKHSLTEIRDSLIDLGLGVFVVPALITARYHRDAGAPVYFE 441
Db      355 --PALADLLMDERTGNSNEBPKHIAQOEWMADAMFVMPALRVAILQR--SHAFTYFE 411
Qy      442 RHRPQCEDTKPAFVKADHDEVAFVFGAFLKGDIVMEFGATEEESKLSSRKMKWTYA 501
Db      412 QHRPSFTKDRPRPHVRADHGEVVFVFRSHLFGSKVPL-----TBEESLLSRRWKTYAN 467
Qy      502 ARTGNPVGNDLSLMPAYNLTEQYIQLDINSLGRLKEPRVDWTSTIP 550
Db      468 ARRNPNVGEGLIAMPFLDQRYQLQLMOPAVGQALKARRLQFWTHTLP 516

```

```

RESULT 3
JC5408
carboxylesterase (EC 3.1.1.1) - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: JC5408
R/Schweizer, H.; Langmann, T.; Daig, R.; Becker, A.; Aslanidis, C.; Schmitz, G.
Biochem. Biophys. Res. Commun. 233, 117-120, 1997
A/Title: Molecular cloning and characterization of a novel putative carboxylesterase, ptc
A/Reference number: JC5408, MUID:97289502, PMID:9144407
A/Accession: JC5408
A/Molecule type: mRNA
A/Residues: 1-559 <SCH>
A/Cross-references: UNIPROT:O00748; GB:Y09616; NID:G2058317; PIDN:CAA70831.1; PID:G205831
A/Experimental source: intestine
C/Comment: This enzyme hydrolyses many xenobiotics, such as carboxyl esters, chloesters ;
C/Genetics:
A/Gene: GDB:CE82; tCE; CE2
A/Cross-references: GDB:9959011
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; glycoprotein
F.158-544/Domain: cholinesterase homology <CH>
F.15-95,123-280,291-428/Disulfide bonds: #status predicted
F.111,276/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.228,457/Active site: Ser, His #status predicted

```

```

Query Match      39.9%; Score 1230; DB 1; Length 559;
Best Local Similarity 46.7%; Pred. No. 1e-88;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

```

```

Qy      31 GPSABGQRMTRIGMIGKQYTVLGSVPVNVFLGPPAPPLGSLRFTNDOPASPMNDL 90
Db      26 GQDSASIRITHTTGQVGLSVHKGANAGVQTLGIPFAPPLGPRFAPPEPESMSG 85
Qy      91 REATSYPNLCLQN-----SEMLLDQMLKMHYPKFVSEDCVLTNLYAPAHDTGSKLP 145
Db      86 RDGTHPAMCLQDLTLVSEFLL-----SQFMTPRDSMEDCXYLSITTPAHSHESNIP 141
Qy      146 VLVWPFQGAFTKTSASIFDGSALAAVEDLVVVVQYRLGIFGFTTWDQAPGNMAFKDQ 205
Db      142 VVWVHGGALVFGMAVLYGSMALALENVVVITIQYRLGIFGFTTWDQAHGHTDQVALLRWYQ 201
Qy      206 VVALSWQKNIIEFGDPSVITFGESAGALSVSSLSLSPMAKGLFKHAIMESGVAILPY 265
Db      202 VALLRWVQGNIAHFGNPNRVITLFGESAGTSVSVLSVPSISGLFPHGAIMESGVALLLG 261
Qy      266 LEADHYEKSBDL--QVNAHFCGNMADSSEALRLCLRTKPSKELTLISQTKSFTRVVDGA 323
Db      262 LIA-----SSADVISTVNAHLSACDDQVDSBALVGCRLGKSKKEILAINKPFKMLPGVVDSV 317
Qy      324 FFPNEBLDLSQAKAPALPSITGVNNHCEGFLP-----MKEAPEILSGNSKLALHLION 379
Db      318 FLDRHPOELASADFPVPSIVGVNNNEFGWLIPIKMYLYDQKMDREASQALQKMLT 377
Qy      380 IHIPOYLHLVANEYFHDKHSLTEIRDSLIDLGLGVFVVPALITARYHRDAGAPVYFE 439
Db      378 LMLLPFTFGDLIREYIIGNDPOQLQAOFOEMADMSFVITALQVAILR--QCSRAPVIFY 436
Qy      440 EPRHAPQCEDTKPAFVKADHDEVAFVFGAFLKGDIVMEFGATEEESKLSSRKMKTYA 499
Db      437 EFQHQPSMLKIRPMPKADHDELPFVF--RSFPGANIKF---TBEEOQLSRKMKWTYA 492
Qy      500 TPARTGNPVGNDLSLMPAYNLTEQYIQLDINSLGRLKEPRVDWTSTIP 550
Db      493 NPARGNPNVGEGLPMPVFLDQEOYQLQLMOPAVGARLKAHRLQFWKRALP 543

```

```

RESULT 4
S47655
carboxylesterase (EC 3.1.1.1) precursor - golden hamster
C/Species: Mesocricetus auratus (golden hamster)
C/Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C/Accession: S47655

```

R:Stone, T.; Isobe, M.; Takabatake, E.; Mang, C.Y.
Biochim. Biophys. Acta 1207, 138-142, 1994
A>Title: Cloning and sequence analysis of a hamster liver cDNA encoding a novel putative
A:Accession number: S47655; MUID:94318665; PMID:8043605
A:Accession: S47655
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <SON>
C:Cross-references: UNIPROT:Q64419; EMBL:D28566; NID:G531238; PTDN:BA05913.1; PID:G531238
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:58-546/Domain: cholinesterase homology <CHE>
F:127,459/Active site: Ser, His #status predicted

Query Match 38.9%; Score 1199; DB 2; Length 561;
Best Local Similarity 47.3%; Pred. No. 2,86-86;
Matches 252; Conservative 74; Mismatches 181; Indels 26; Gaps 7;

D Db 31 GPSAEGRNRRLGWIGKQVTVLGSPVPVVNLGPAAPPLGSIAFTNPQAPSPPDNL 90
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
26 QGDVSPIRNTHTQVRGKVLYYREGVTGYAFGLGPAPRPVGLPFAPEEPPEPSGV 85
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 91 REATSYENLCLONSEMTLLDDHMLKVHPKPGVSEDCIYNIVYPADAGSKLPVLWVF 150
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 86 RDGISEBAMCLQDPMPAPOLSKERRILLPTISBEDCLINITYPALAHESNLPVAVMI 145
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 151 PGCAFKTGSASIFDGSLAAAYEDLVVVVVOYRGLGFSEFTTMDQAGANNAFKQVAALS 210
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 146 HGGALVMGMASMNDSILLAATEDIVVISIQYRLGILGFSTGDEHARGNWGYLDQVALH 205
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 211 WVGKNIEFPFGDBSVITTFGESAGAISSLSILSPMAKGLPHKAIMSGVALIIFYLEAND 270
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 206 WVQONIASFGSNPQQVITIFGVSAQTSSVSLVSPMSKGLFHGAIMOSGVALLPDL-1SD 264
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 271 YEKSEDQVAAFPCGNASDEALLRCIRTPSKELLTLISQTKSPTFRVVDGAFEPNEPL 330
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 265 TPBALVPRVANNGSCENKSEALVHCIREKTEAILAINOVFMETGVGDGFLPRHPQ 324
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 331 DLTSQAKFKAIPSIIGVNNHECGFLPM-----KEAPEILSGNSKLALHLIONIL 381
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 325 ELIASVDFHPVPISIIIGVDSDECGVPLPFMGLDHVINKITRETLPAVLKSABHHM---- 380
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 382 HIPPOYLVLANEFFDHSHLTETRDSDLGLGVFVVPALITARYHRDAQAVTYFER 441
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 381 -LPRECDDLMOEFTMGVEDVEPOTLQAOFRELMDKFVFIIPALKAVAFQR-SHAVVYYEF 438
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 442 RHRQCC--EDTKAFPKADHADDEVRFVGCFAF--LKGDIYMFGCAFEERKLSRKMKY 497
| : | : | : | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 439 QHOSSFINKDKARSHRADGHDAVAFVSGDFGLKIDL-----TEEBELTLKRMMKY 492
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 498 WATGARTGNPNGNDSLMPAYNLTEOYLQLDNLSLGORLKPEPVDFWTSTIP 550
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 493 WANARRGNRNSEGLPYMPBELVHDQYLKDIOAVGRALKSRGLHWTTILP 545
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
S34607
carboxylesterase (EC 3.1.1.1) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S34607
A>Title: Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase 1
C:Biochim. Biophys. Acta 1174, 72-74, 1993
R:Aida, K.; Moore, R.; Negishi, M.
A:Reference number: S34607; MUID:93326638; PMID:7916639
A:Accession: S34607
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <AID>
C:Cross-references: UNIPROT:Q63880; GB:S64130; NID:G404388; PTDN:AA027606.1; PID:G404389
C:Keywords: carboxylic ester hydrolase
F:46-536/Domain: cholinesterase homology <CHE>

[illegible]

```

Db      11 LAVCPILGH-----SLLPVVDITGKVLGKYSLSGFEQPVAVLGVPPAKPPLGS 62
Qy      76 LRFNTPOPASPMWNLREATSYPNLCLONSEM--LLDDHMLKXHYKFGVSEDCLYLNTY 133
Db      63 LRFAPBPAPBPSFVKNTATSYPPMCSQDAGMAKILSDMFSTEKEILLPLKISEDCLYLNTY 122
Qy      134 APAAHDTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVQYRLGIFGFFTTMD 193
Db      123 SPADLTKRSQLPVWVWIGHGGGLVIGGRSPYNGALSAHENVVYTIQYHLGIGLGFSTGD 182
Qy      194 QHAPGMAKQDVAALSWQKNIIEFFGCDPSSVTTFGESAGAISSVSLILSPMAKGLERK 253
Db      183 EHSPPGMALDQDLALRWQDNIANFGGPDSTVTTFGESGGSIVSLVLSPLGKDLFHR 242
Qy      254 AIMESGVAIIPYLEAHDYKESDLOVVAHFCGNMNSDSEALLRCITRKSKLLTLS--- 310
Db      243 AISESGVAVNTAVGKKNIOAVNEIATLSQC--NDTSSAAWOCROKTESSEILISGKL 300
Qy      311 -OKTYSFTRVVDGAFFPNBPLDLLSQAKFKAIPSIIGVNNHCGFLLPM---KEAPE-IL 365
Db      301 VQYNISLSTMIDGVLPKAPBEILAEKSPNTVPYIVGFVKQGFMIIPMQLONLLPEGKM 360
Qy      366 SGSNKSILHLQNLHTIPQYLHVANEFYHDKSLTEIRDSLLDLDGDFVFPVALIT 425
Db      361 NBETASLLRRFHSLENTSESMTIPVIEQYLGVDDPAKKSLLIDMFGDIFFGIIPAVIL 420
Qy      426 ARYHDAQAPVYEFYERHRRPOCFEDTKPAFVADHAEVRFVFGGAPLKGDIVMFEGATE 485
Db      421 SRSLDAGVSTVWYFRFRPSFVSDKRPGTVBGDGHDELFFVFGAPLAK-----EGASE 474
Qy      466 BEKLSRKMKKWTAPARTGNPNNDLSLWPAVNLTEOYLQDLNLSGRLKEBRYDFW 545
Db      475 EETNLSKVMKMANFARNGNFNGELPHWPEYDQEGYLQIGATTQOAGRLKAEVAFW 534
Qy      546 TSTI 549
Db      535 TELL 538

```

RESULT 7

```

S19307
C:carboxylesterase (EC 3.1.1.1) precursor - pig
N:Alternate names: proline-beta-naphthylamidase
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S19307; S23607
R:Matsumura, M.; Inoue, H.; Ichinose, M.; Teukada, S.; Miki, K.; Kurokawa, K.; Takahashi, T. 1991
P:FEBS Lett. 293, 37-41, 1991
A:Title: The nucleotide and deduced amino acid sequences of porcine liver proline-beta-naphthylamidase
A:Reference number: S19307; MUID:92070571; PMID:11959668
A:Accession: S19307
A:Molecule type: mRNA
A:Residues: 1-566 <MAT1>
A:Cross-references: UNIPROT:Q29550; EMBL:X63323; NID:g1930; PIDN:CAA44929.1; PID:g1931
A:Accession: S23607
A:Molecule type: protein
A:Residues: 19-40 <MAT2>
A>Note: 28-Lys and 33-Leu were also found
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-566/Product: carboxylesterase #status experimental <MAT>
F:51-552/Domain: cholinesterase homology <CHB>
F:80/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:222,467/Active site: Ser, His #status predicted

```

Query Match 37.8%; Score 1162.5; DB 2; Length 566;

Best Local Similarity 44.3%; Pred. No. 2,1e-83;

Matches 250; Conservative 71; Mismatches 202; Indels 41; Gaps 8;

Qy 12 WCFLLIPLGHRQWKTGPSAEGPQRNRLGNIQKQVTVLGSVPVNVVFLGVFAAP 71
 Db 2 WLLPLVLTSLASATW--AGQPASPPVVDTAQGRVIGKYSLSGLGAPVAVFLGVFAAP 59

```

Qy      72 PLGSLRFTNPQASPMWNLREATSYPNLCLON-----SEWLLDDHMLKXHYKFGVS 124
Db      60 PLGSLRFAPQAPBAPBPSFVKNTTSPMCSQDPAVBQGTSDLPFNGBERLLEB-----S 114
Qy      125 EDCLYLNTYAAHADTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVQYRLG 184
Db      115 EDCLYLNTYTAADLTKGRGLFVWVWIGHGGGLVIGGAPYDGVVLAHENVVVAIQRGL 174
Qy      185 IGFFTTMDQAPGNMAFKQVAALSWQKNIIEFFGCDPSSVTTFGESAGAISSVSLILS 244
Db      175 IMGFFSTGDEHSRGNMGHLQVAAHNVQENIANFGDPSGVTTFGESAGGESVSLVLS 234
Qy      245 PMAKLTHKAIMESGVAIIPYLEAHDYKESDLOVVAHFCGNMNSDSEALLRCITRKPSK 304
Db      235 PLAKLTHFRALISEGVALTVVALVRKDKAA--AKQIVLAGCKTTTSAVVHCLRGQSED 292
Qy      305 ELTLTQKTSFT-----RVDDGAFPPNEPLDLLSQAFKAIPSIIGVN 349
Db      293 ELDDLTKMKFTLIDFGDQRESHPLPVTVDGVLKPMKEEILAEKDFNTVPYVIGINK 352
Qy      350 HECGFLLPMKREPIILSGS-NKSLALHLQNL---LHIPPQYLHVANEFYHDKSLTEI 405
Db      353 QEFGMLLPTMMGFPLSGKLDQKATATSLWKSYPINIPBELTPVATDKVLGGTDPPVK 412
Qy      406 RDSLLDLDGDFVFPVALITARVHRDAGAPVYFERRHRRPOCFEDTKPAFVADHAEVR 465
Db      413 KDLFLDLMGVVFGVPSVTVARQRDAGAPTYVEFOYRFSFSSDKPKPTVIGDHDDEIF 472
Qy      466 FVFGAFLKGDIVMFEGATEEEKLSRKMKKYATPARTGNPNNDLSLWPAVNLTEOYL 525
Db      473 SVFGFPLTKD-----APBEVSLSKTYWKPFANFARSGNPNGEGLPHWPEYDQEGYL 526
Qy      526 QDLNLSGRLKEBRYDFWFTSTI 549
Db      527 QIGVNTQAAKRLKGEVAFWMDLL 550

```

RESULT 8

```

JX0054
C:carboxylesterase (EC 3.1.1.1) El precursor, minor form - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Aug-1998
C:Accession: JX0054
R:Takegi, Y.; Morohashi, K.; Kawabata, S.; Go, M.; Omura, T.
J. Biochem. 104, 801-806, 1988
A:Title: Molecular cloning and nucleotide sequence of cDNA of microsomal carboxylesterase
A:Reference number: JX0054; MUID:89174514; PMID:3235453
A:Accession: JX0054
A:Molecule type: mRNA
A:Residues: 1-549 <TRX>
A:Experimental source: liver
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein; microsome
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-549/Product: carboxylesterase El #status predicted <MAT>
F:50-538/Domain: cholinesterase homology <CHB>
F:79,274,302,375,476/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:221,453/Active site: Ser, His #status predicted

```

Query Match 37.2%; Score 1144; DB 2; Length 549;

Best Local Similarity 43.5%; Pred. No. 5,8e-82;

Matches 238; Conservative 92; Mismatches 183; Indels 34; Gaps 9;

Qy 16 LILPLPLGHRQWKTGPSAEGPQRNRLGNIQKQVTVLGSVPVNVVFLGVFAAPPLGS 75
 Db 11 LAVCPILGH-----PSSP-PVVDITGKVLGKYSLSGFTQPAVFLGVFAAPPLGS 62
 Qy 76 LRFNTPOPASPMWNLREATSYPNLCLON-----SEWLLDDHMLKXHYKFGVSEDC 128
 Db 63 LRFAPBPAPBPSFVKNTTTPMCSQDQGVGKLIADLSTGKSIEPLF-----SEDC 117
 Qy 129 YLNTYAAHADTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVQYRLGIFG 188


```

Db 118 YLNIYSPADLTKNRSLPVMWVHGGGLIGGASPYSGALSAHEVVVVTIQLRGIWGL 177
Qy 189 FTTMDQAPGMAKRDQVAAALSMVQKNIIEFGDPSSVTITIGESAGAISVSLLISPMAX 248
Db 178 FSTDEHSRGMWALHDOALALRWODNIANFGGNDSTIIGESAGVSVLALYSPILAK 237
Qy 249 GLPKAIMEGVAIIIPYLEADHYEKSEDLQVVAHCGNNADEBALRCLRTPKSKELLT 308
Db 238 NLFRALISEGVALTLTNLDKQTOAV--AQMIATISGNTNTSSAAMVOCLOKTEALLE 295
Qy 309 LSQK-TKSPTRVVDGAFFPNEPLDLSQKAFKAIPIIIGVNNHCEGLPMKCAPEILS 366
Db 296 LTVGLDNTSMSTVIDGVLPKTPBEILTEKSFNTPVYIVGNKQEFWIIPTMGNLILSE 355
Qy 367 G-SKSLALHLIQLI--LHIPPQYLIVANEFYFDKSLTEIRDSLIDLGLDVPFVPA 422
Db 356 GRMKEKMASSFLKRPSPMLNTSESVTPAIIEKYLRTDPAKKKELLDMESDVFFGIPA 415
Qy 423 LITARYHRDAGAPYFYEFRFRPOCFEDTKPAFYKADHADVRFVFGGAFPKGDIVMPEG 482
Db 416 VLMSRLRDAGAPYFTEFQYRPSVSDQRPQYQGDHGDIFSVYFGPIFK-----EG 469
Qy 483 ATBEKSLSRKMKYMATPARTGNPNGLDSLMPAYNLTEOYLQDLNMSLGORLKEPRV 542
Db 470 ASBETMLSKLVKMFMANFANGNPNBGLPHWPEYDQKEGYLQIGATTQQAQKLKGEV 529
Qy 543 DFWTSTI 549
Db 530 AFWTELL 536

```

RESULT 9

```

A41010
carboxylesterase (EC 3.1.1.1) precursor, monocyte/macrophage [validated] - human
N:Alternate names: carboxylesterase, hepatic; monocyte/macrophage serine esterase
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C:Accession: A41010, JH0327, A47376, A49816, PS0280, I61085, A48809, I57004
R:Munger, J.S.; Shi, G.P.; Marx, E.A.; Chin, D.T.; Gerard, C.; Chapman, H.A.
J. Biol. Chem. 266, 18832-18838, 1991
A:Title: A serine esterase released by human alveolar macrophages is closely related to
A:Reference number: A41010; MUID:92011649; PMID:1918003
A:Accession: A41010
A:Molecule type: mRNA
A:Residues: 1-567 <MON>
A:Cross-references: UNIPROT:P23141; GB:M73499; NID:9179927; PID:AAA35649.1; PID:9179928
A:Note: parts of this sequence, including the amino end of the mature protein, were conf
R:Long, R.M.; Calabrese, M.R.; Martin, B.M.; Pohl, L.R.
Life Sci. 48, PL43-PL49, 1991
A:Title: Cloning and sequencing of a human liver carboxylesterase isoenzyme.
A:Reference number: JH0327; MUID:91148424; PMID:1997784
A:Accession: JH0327
A:Molecule type: mRNA
A:Residues: 61-567 <ION>
A:Cross-references: GB:M55509; NID:9179929; PID:AAA35650.1; PID:9179930
R:Shibata, F.; Takagi, Y.; Kikajima, M.; Kuroda, T.; Omura, T.
Genomics 17, 76-82, 1993
A:Title: Molecular cloning and characterization of a human carboxylesterase gene.
A:Reference number: A47376; MUID:94010913; PMID:8406473
A:Accession: A47376
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA; DNA
A:Residues: 1-3, 'PALV', 8-11, 'A', 13-567 <SHI>
A:Cross-references: GB:D21088; NID:945476; PID:BA04650.1; PID:9458470
A:Note: sequence extracted from NCBI backbone (NCBI:P137630) and corrected to correspond
R:Zschunke, F.; Salmasi, A.; Kreipe, H.; Buck, F.; Parwaresch, M.R.; Radzun, H.J.
Blood 78, 506-512, 1991
A:Title: cDNA cloning and characterization of human monocyte/macrophage serine esterase-
A:Accession: A49816; MUID:91300111; PMID:2070086
A:Molecule type: mRNA
A:Residues: 'G', 65-185, 'G', 187-361, 363-567 <ZSC>

```

```

A:Cross-references: GB:X52973; NID:936421; PID:CAA37147.1; PID:91335304
R:Riddles, P.W.; Richard, L.J.; Bowles, M.R.; Pond, S.M.
Gene 108, 289-292, 1991
A:Title: Cloning and analysis of a cDNA encoding a human liver carboxylesterase.
A:Reference number: PS0280; MUID:92084150; PMID:1748313
A:Accession: PS0280
A:Molecule type: mRNA
A:Residues: 114, 'R', 116-280, 'A', 282-300, 'IGNSLWYRETOREST', 318-336, 'R', 338-382, 'GSP', 384
A:Cross-references: GB:M65261; NID:9187028; PID:AAA83932.1; PID:9187029
A:Experimental source: liver
A:Note: differences between this sequence and other reports appear to be due to frameshift
R:Kroetz, D.L.; McBride, O.W.; Gonzalez, F.J.
Biochemistry 32, 11606-11617, 1993
A:Title: Glycosylation-dependent activity of baculovirus-expressed human liver carboxyle
A:Reference number: A48809; MUID:94032283; PMID:8218228
A:Accession: I61085
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-55, 'G', 57-361, 363-535, 'G', 537-567 <KRO1>
A:Cross-references: GB:I07765; NID:9180949; PID:AAA5711.1; PID:9180950
A:Accession: A48809
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-17, 'A', 18-55, 'G', 57-535, 'G', 537-567 <KRO2>
A:Cross-references: GB:I07764; NID:9180947; PID:AAA16036.1; PID:9180948
A:Genetics:
A:Gene: GDB:CEB1, HMSR
A:Cross-references: GDB:128044, OMIM:114835
A:Map position: 16q13-16q22.1
A:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; liver
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-145/Domain: propeptide #status predicted <PRO>
F:150-553/Domain: cholinesterase homology <CHB>
F:146-567/Product: carboxylesterase #status experimental <MAT>
F:564-567/Region: endoplasmic reticulum retention signal #status atypical
F:221,468/Active site: Ser, His #status predicted
Query Match 37.0%; Score 1140; DB 1; Length 567;
Best Local Similarity 44.3%; Pred. No. 1.3e-81;
Matches 248; Conservative 79; Mismatches 197; Indels 36; Gaps 9;
Db 12 WCFPLIPLGLGHQMKCTGPSAEGPQNTLGIQOKOYTVLGSPPVNVFLGVPAP 71
2 WLRAPILATLSASAKWH--PSSP-PVVDYTHGKVLKFPYSLBGFAPVAFILIPFAKP 58
Qy 72 PLGSLRFTNPQPAPSPMDNLEATSYPNLCLQNSE--WLLDQHLKVVHYKGVSEDCLY 129
Db 59 PLGRLRFTPPQAPAPMSEFVKATSYPPMCQDPKAGQLSELFTNRKENIPLKLESDCLY 118
Qy 130 LNTAPPAHADTSGCLPVLWFPFGCAFRTGASIFDGSALAAVEVNVVVOYRIGIGCF 189
Db 119 LNTYTPADLTCKNKLPMWVWVHGGGLWGAASYDGLALAAHEVVVVVTIQLRGIWGF 178
Qy 190 TTMDOAPGMAKRDQVAAALSMVQKNIIEFGDPSSVTITIGESAGAISVSLLISPMAX 249
Db 179 STGBEHSRGMWALHDOALALRWODNIANFGGNDSTIIGESAGVSVLALYSPILAK 238
Qy 250 LFRKAIMESGVAIIIPYLEADHYEK--SEDLQVVAHCGNNADEBALRCLRTPKSKELL 307
Db 239 LFRHAISESGVALTSVLYKKGDVFLBQALITA---GCKTTTSAMVWHCLRQTEBEL 295
Qy 308 TLSQRTKSFY-----RYVDGAFPPNEPLDLSQKAFKAIPIIIGVNNHCE 352
Db 296 ETTIKMFTLSLDQGDRESQPLIGTYIDGLMLKTPBELQAERNFTVPMVGINKQEF 355
Qy 353 GLIPMKCAPEILSGS--NKSIALHLIQLI--LHIPPQYLIVANEFYFDKSLTEIRSD 407
Db 356 GMLIPQQLMSYPLSEGGQLDQKATSLMKSPYLVCIAKELIPATREKTLGTDITVKKD 415
Qy 408 SLDDLDDVPFVVPALITARYHRDAGAPYFYEFRFRPOCFEDTKPAFYKADHADVRFV 467
Db 416 LFLDLIDVMEGVSVIVARHRDAGAPTYMYEFQYRPSFSSDMKPKTVIGDHDELFSV 475

```

D

```

QY      468 FGAFPLKGDIVMEFGALREEEKLSRKMKWMAFPARTGPNPNDGLSLMPAINVLTQYL 527
           ||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db       476 FGAPFLR-----EGASEEIRLRSKWVKWMANFARNGNPGRLPHMPEYNOKEGYLOI 529
                               |::| |::|
QY      528 DLNMSLGQRLEPRVPWFMTS 547
Dd       530 GANTQAARQLKDKEVAFWTN 549

RESULT 10
562788
carboxylesterase (EC 3.1.1.1) ES-4 precursor, liver - rat
N:Alternate names: hydrolase B
C/Species: Rattus norvegicus (Norway rat)
C/date: 28-Oct-1996 #sequence revision 13-Mar-1997 #text_change 09-Jul-2004
C/accession: S62788; SS1203; AS5304; S49257
R/Kobbi, M.; Van Schattingen, E.; Beaufay, H.
Biochem. J. 313, 821-826, 1996
A>Title: Cloning and sequencing of rat liver carboxylesterase ES-4 (microsomal palmitoyl
A/reference number: S62788; MWID:96190723; PMID:6611161
A/accession: S62788
A:molecule type: mRNA
A:Residues: 1-561 <ROB>
A/Cross-references: UNIPROT:O64573; EMBL:X81825; NID:G550417; PIDN:CNA57419.1; PID:G550418
A/experimental source: Liver
R/Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A>Title: Purification and characterization of two rat liver microsomal carboxylesterases
A/accession: S51203; MWID:95077430; PMID:7986098
A/molecule type: protein
A:Residues: 19-48 <MOB>
A/Experimental source: Liver
R/Yan, B.; Yang, D.; Brady, M.; Parkinson, A.
J. Biol. Chem. 269, 29688-29696, 1994
A>Title: Rat kidney carboxylesterase. Cloning, sequencing, cellular localization, and re
A/reference number: AS5304; MWID:95050819; PMID:7961958
A/accession: AS5304
A:molecule type: mRNA
A:Residues: 1-6,'F','R'-59,'P','61-212,'A','214-252,'T','254-309,'IT','312-341,'N','343-424,'FY
F1:Cross-references: GB-U010697; NID:G562007
R/Rocchi, M.; Beaufay, H.
submitted to the EMBL Data Library, September 1994
A/reference number: S49257
A/accession: S49257
A/statute: preliminary
A:molecule type: mRNA
A:Residues: 1-561 <RO>
A/Cross-references: EMBL:X81825; NID:G550417; PIDN:CNA57419.1; PID:G550418
C/function:
A:description: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
A/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolysis; endoplasmic reticulum; glycoprotein; liver
F:1-18/Domin: signal sequence #statue predicted <IG>
F:19-561/Product: carboxylesterase ES-4 #statue experimental <MAT>
F:150-551/Domian: cholinesterase homology <CHB>
F:221,466/Active site: Ser, His #statue predicted
F:301/Binding site: carbohydrate (Asn) (covalent) #statue predicted

Query Match          37.0% Score 1139; DB 2; Length 561;
Best Local Similarity 43.4%; Pred. NO. 1.Se-81;
Matches 243; Conservative 82; Mismatches 195; Indels 40; Gaps 10;

QY      15 FLILQPLLGRWGKTGPSAEQPQNTRLCMIQGKVTVLGSFVNPFAGVFAPAPIG 74
           ||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db       5 FLIIVSLATCVVG--NPSSP-PVVDTKKKVLGVKSVELEGYSAVNLGCPFAEPPIG 61

QY      75 SLSEFTNQPSPPMDNRKATSYRNLCIONSEM-----LLIDQMMLNVHYPKFVSDCLY 129
           ||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db       62 SLSEFAPQPEEPSFVNNTTTTPPMCSODAKQGMNDDLITNRKEKH---LEFSEDCLY 118
    
```

[illegible]

RESULT 11

carboxylesterase (EC 3.1.1.1) egasyn - mouse
N/Alternate names: beta-glucuronidase endoplasmaic reticulum-targeting protein; esterase-2
C/Species: Mus musculus (house mouse)
C/Date: 10-Feb-1995 #sequence_rev1sion 10-Feb-1995 #ext_change 09-Jul-2004
C/Accession: A55281
R/Ovnc: M.; Swank, R.T.; Fletcher, C.; Zhen, L.; Novak, E.K.; Baumann, H.; Heintz, N.; C
Genomics 11, 956-967, 1991
A/Title: Characterization and functional expression of a cDNA encoding egasyn (esterase-2)
A/Reference number: A55281, MIMD:92147141, PMID:1783403
A/Accession: A55281
A/Status: preliminary
A/Molecule type: mRNA; protein
A/Residues: 1-562 <UNP>
A/Cross-references: UNIPROT:064176; GB:S80191; NID:9244727; PIDD:AA821335.1; PTD:9244728
A/Note: sequence extracted from NCBI backbone (NCBIN:80191, NCBIP:80194)
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; endoplasmaic reticulum
F:S1-552/Domain: cholinesterase homology <CHE>
F:222,467/Active site: Ser, His #status predicted

	37.0%;	Score 1139.5;	DB 2;	Length 562;
Query Match	Best Local Similarity	44.4%;	Pred. No. 1.6e-81;	
Matches 240;	Conservative	76;	Mismatches 195;	Indels 29; Gaps 6
QY	31	GSASAGPQQRNRLGHIQKQVTVLGSFVFNVEFGVPPAAPPLGSLRFTTNPQDASPMIDL	90	
DB	19	GHSPSSPPWVDVVGQGVNIGKXISLSGFTQPAVFLGVFPAPPLGSLRFPAPQPAEPMSSV	78	
QY	91	REATSYPNLICLQNSEWMLLDQHMILKVAHPKFGV--SEDCYLNIYAPAHADTSSKLPVLY	148	
DB	79	KNATSYPMPCQDVPVQGIIVNDLITNRKKEKPIQFSFDCILYNIYTPADULTSDRLPPVY	138	
QY	149	WPPGSAFPTGASLIPDGSALAAVEDVLYVVVQVRLGIFGFTTWDQAHGNAAFKQVAA	208	
DB	139	MIHGGLVLTGASATYDGLVLTSHENVVVVVYQRLGIGWFSFGSDHSKRNMGHLDQVAA	198	
QY	209	LSWVQKNIFFGQDPSSVTTFGSSAGAISSVSLISPMAKGLFHKAIMESGVAIIPYLEA	268	

199 LHMWDNIATKPGDGSVTTIFGESAGGSVSVLVISPLAKNLFORALISSEGVALLTAGLVK 258
 269 HDYKSEDLQVVAHFCCGNASDSEALLRCLRTKPSKELLTTSQKRSF----- 316
 259 KOTRPLAEKIAVIGCCKN--TTSAAVHCLROKTEEBELGTLKNTFLDLHDSRQSH 316
 317 ---TRVDGAFPPNEPLDLSQAKFAIPSLIGVNNHCEGFLP--MKEAPEILSGSNKSL 372
 317 PFVPTLVLDGVALPKKBEELIAEKKNFTVPIYVINKQSGFWLPTMMVYPPEDVLDQMT 376
 373 ALHLIQN---LHLPPOYLHLVANEYFHDKSLTEIRDSLLDLGDVFFVPALITARYH 429
 377 AMSLLKSSFLPNTPEDAIAVAIEKRLDKOYTGKNGQQLLELDVVGVPSPVIVSGH 436
 430 RDAGAPVYFYEFRHPPQCFEDTKPAFVKADHDEVFPVFGCAFLLGDLVMEFGATEEEL 489
 437 RDAGAPVYFYEFRHPPQCFEDTKPAFVKADHDEVFPVFGCAFLLGDLVMEFGATEEEL 490
 490 LSRKMKYATPARTGNGNDLSMPVYNLTQYLQDLNMSLQRLKEPPVDEWTSTI 549
 491 LSRKMKYATPARTGNGNDLSMPVYNLTQYLQDLNMSLQRLKEPPVDEWTSTI 550

RESULT 12

A:1584
 A:Accession: A1584
 A:Residues: 1-540 <UNP>
 A:Cross-references: UNIPROT:P10599; GB:M20629; GB:X13587; NID:G203279; PIDN:AAA0871.1;
 C:Keywords: carboxylic ester hydrolase; cholinesterase homology
 F:1-9/Domain: signal sequence #status predicted <SIG>
 F:10-540/Product: carboxylesterase #status predicted <MAT>
 F:41-529/Domains: cholinesterase homology <CH>
 F:70,265,266,293,366,467/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:212,444/Active site: Ser, His #status predicted
 Query Match 36.9%; Score 1136; DB 2; Length 540;
 Best Local Similarity 43.3%; Pred. No. 2,4e-81;
 Matches 237; Conservative 91; Mismatches 185; Indels 34; Gaps 9;

16 LILQPLIGHRQMGKTSABGQRNTRIGWIOGQVTVLGSVPVNVFLGVPAPAPPLGS 75
 2 LAVCPIMWH-----PSSP--PVVDITKGVKLVKVSLEGFTQPAVFLGVFPAPPLGS 53
 76 LRFNTPOPASPMNDLREATSYNLTQNL-----SEWLLDQHLKHYPRKGVSEDL 128
 54 LRFAPPEABEPMSPKNTTTPKMGSDGVGKLAIDLSTGKESIPLEF-----SEDL 108
 129 YLNIYAPPAHADTGSCLPVLVMPFGAFKTSASTIDGSALAAVEDVLYVVOYRIGIFG 188
 109 YLNIYSPADLTNKSRLPVMWVHGGGLIIGGASPYSGALSAHEVNVVVTTIQYRIGFGL 168
 189 FTTMDQAHGMMAFDOVAALSMVOKNIEFPGDSSVTIFGESAGTSSVSLISPMAX 248
 169 FSTDEHSRGMWMLDQALALRWODNIANFGANDSVTIFGESAGTSSVSLISPLAK 228
 249 GLFKAIMESGVAIIPYLEADHYEKSEDLQVVAHFCCGNASDSEALLRCLRTKPSKELLT 308
 229 NLFRALISESGVVLITNDKKNQAV--AQMIATISGCGNTSSAAMVOCILKTEABELLE 286
 309 LSQK--TKSFRVVDGAFPPNEPLDLSQAKFAIPSLIGVNNHCEGFLP--MKEAPEILSG 366
 287 LTVLDNTSMSTVLDGVVLPKPBELITEKSFNTVPIYVINKQSGFWLPTMMVYPPEDVLDQMT 346

367 G-SNKSIALHLIQN---LHLPPOYLHLVANEYFHDKSLTEIRDSLLDLGDVFFVPA 422
 347 GMMNKMASSFLKRSRPNLINESVYIPALIEKLYNGTDPACKKELLDMFSDVFGIARA 406
 423 LITARYHDDAGAPVYFYEFRHPPQCFEDTKPAFVKADHDEVFPVFGCAFLLGDLVMEFG 482
 407 VLMSRLSDAGAPVYFYEFRHPPQCFEDTKPAFVKADHDEVFPVFGCAFLLGDLVMEFG 460
 483 ATEEBKLSRKMKYATPARTGNGNDLSMPVYNLTQYLQDLNMSLQRLKEPPVDEWTSTI 542
 461 ASEEBKLSRKMKYATPARTGNGNDLSMPVYNLTQYLQDLNMSLQRLKEPPVDEWTSTI 520
 543 DFWTSTI 549
 521 AFWTELL 527

RESULT 13

A:1597
 A:Accession: S71597
 A:Residues: 1-561 <YAN>
 A:Experimental source: liver; endoplasmic reticulum
 A:Function: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
 C:Superfamily: cholinesterase; cholinesterase homology
 C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-561/Product: carboxylesterase #status predicted <MAT>
 F:50-551/Domains: cholinesterase homology <CH>
 F:558-561/Region: endoplasmic reticulum retention signal
 F:79,301/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:221,466/Active site: Ser, His #status predicted
 Query Match 36.7%; Score 1129.5; DB 2; Length 561;
 Best Local Similarity 43.4%; Pred. No. 8.3e-81;
 Matches 236; Conservative 83; Mismatches 188; Indels 37; Gaps 9;

31 GPSABGPQRNTRIGWIOGQVTVLGSVPVNVFLGVPAPAPPLGSIRFTNPOASPMNDL 90
 18 GNPSSPPVVDITKGVKLVKVSLEGFTQPAVFLGVFPAPAPPLGSIRFTNPOASPMNDL 77
 91 REATSYNLTQNLSEW-----LLDQHLKHYPRKGVSEDLTYLNIYAPPAHADTGSCLP 145
 78 KNTTTPKMGSDGVGKLAIDLSTGKESIPLEF-----SEDL 134
 146 VLVMPFGAFKTSASTIDGSALAAVEDVLYVVOYRIGIFGFTTMDQAHGMMAFDO 205
 135 VVMWVHGGGLIIGGASPYSGALSAHEVNVVVTTIQYRIGIFGFTTMDQAHGMMAFDO 194
 206 VVALSMVOKNIEFPGDSSVTIFGESAGTSSVSLISPMAXGLFKAIMESGVAIIPY 265
 195 VVALHMDNIANFGANDSVTIFGESAGTSSVSLISPMAXGLFKAIMESGVAIIPY 254
 266 LEADHYEKSEDLQVVAHFCCGNASDSEALLRCLRTKPSKELLTTSQKRSF-----TK 314
 255 LFTKQVRA--AKQADWAGCKTTTSAIIVHCLROKTEEBELIEKKNLILKSSORDTK 312
 315 S-----FTVVDGAFPPNEPLDLSQAKFAIPSLIGVNNHCEGFLP--MKEAPEILSG 368
 313 ESYHFLSTVLDGVVLPKPBELITEKSFNTVPIYVINKQSGFWLPTMMVYPPEDVLDQMT 371
 369 NKSIALHLIQ---LHLPPOYLHLVANEYFHDKSLTEIRDSLLDLGDVFFVPA 425

```

Db      372 DKQAMTLEKFASTIGIPEDIIIVPAIEKRRKSDDEPIKIRDRDILAFIGDVFICISVWV 431
Qy      426 ARHRRAGAPVYTYEERHRRQCEBDTKPAFVKADHADEVRFVGAFELKGDIMFEGATE 485
Db      432 SRHRAGAPVYTYEYQYPSFSSPQRPKDVGDHADDVYSVGAIDLK-----DGASE 485
Qy      486 EEKLSRKMMKYATAPARTGNPNNDLSLMPAYNLAEVYLQDLNMSLQORLKEPRVDW 545
Db      486 EEKLSRKMMKYATAPARTGNPNNDLSLMPAYNLAEVYLQDLNMSLQORLKEPRVDW 545
Qy      546 TSTI 549
Db      546 TOLL 549

RESULT 14
JC2447
carboxylesterase (EC 3.1.1.1) ES-3 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Mar-1995 #sequence _revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: JC2447; S23462
R:Robbi, M.; Beaufay, H.
Biochem. Biophys. Res. Commun. 203, 1404-1411, 1994
A:Title: Cloning and sequencing of rat liver carboxylesterase ES-3 (BGASYN).
A:Reference number: JC2447; MUID:95032008; PMID:7945287
A:Accession: JC2447
A:Molecule type: mRNA
A:Residues: 1-561 <ROB>
A:Cross-references: UNIPROT:063108; GB:X81395; NID:9550146; PIDN:CAA57158.1; PID:9550147
A:Experimental source: liver
R:Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A:Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A:Reference number: S23460; MUID:92299008; PMID:1606962
A:Accession: S23462
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 503-554, 'R', 556-561 <MED>
A:Cross-references: EMBL:X65295; NID:957557; PIDN:CAA46390.1; PID:957558
C:Superfamily: Cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-551/Product: carboxylesterase ES-3 #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CHES>
F:79,107,489/Binding site: carbonyl group (Asn) (covalent) #status predicted
F:221,466/Active site: Ser, His #status predicted

Query Match 36.5%; Score 1123; DB 2; Length 561;
Best Local Similarity 42.2%; Pred. No. 2.7e-80;
Matches 245; Conservative 86; Mismatches 193; Indels 56; Gaps 11;

Qy      13 CFFLLDPLRGHROWGKTGSAEGFORNTRLGWIOGKQVTVLGSPPVNVVLFVGPAPR 72
Db      2 CLVALILVFLAFTAG--GHPSSLPVVDTLQGVKLVLEGTQVAVVFLGVPKAPR 59
Qy      73 LGSIRFTNPQASAWNDLREKTSIPNLCLO-----SEWLLIDHMLKVHYPKF 121
Db      60 LGSIRFAPPOPAEAWSPVKNTSTYPMCSODPVAGQIVNDLITW-----EENISLQF--- 112
Qy      122 GVSEDCVLYNLYAPAHADTGSKLPLVWFPQGAFTKTSASIFDSALAAEDVLYVVVQY 181
Db      113 --SEDCVLYNLYTADTKDRPLVWVWVHGGGLVLDGASITVDGLSTHENVVVVVQY 170
Qy      182 RLGIFFGFTTWDQAPGNMAFKQVVALSWYQKNIEFFGDPSSVTTFGSSAGAISSVSL 241
Db      171 RLGIWFGFTSGDEHSRGNMGHLQVVALHWQDINDFGDDPGSVTLFGSSAGSESVSL 230
Qy      242 ILSPMAKGLHKKALMESGVALLPYL--BAHYEKSEDDQVVAHCGNNAASELLRLCLRT 300
Db      231 VLSPLANLHKKALSESGLVALTGLVKKNTRPLAEKLAIVS---GCSTTSASMVHCLRQ 287
Qy      301 KPSKELLTLISQKTSFT-----RVVDGAFPPNBPDLDSGAKAKALPSII 345

```

```

Db      288 KTEEBELTTTKLMLFSLDLDHDSRQSPVFPVTVLDGVVLKMEBEILAEKDFNTVPYIV 347
Qy      346 GVNNNECGFLP--MKEAPEILSGSKSLAHLION---ILHIPQVYLVAENVFHDKS 401
Db      348 GINKEPFWIILPTMANNYPSPDKLDPMTATISLKKSSFLMLPBEALPVAVEKTLRHTDD 407
Qy      402 LTERDLSLIDLIGVFFVPPALITARYHRDAGAVTYEERHRRQCEBDTKPAFVKADHA 461
Db      408 PDRNKDQLLEILIGDIVIFGVPSVIVSRGHRDAGARTYMEFQYRPSFSSKKPSTVGDHG 467
Qy      462 DEAFVFGGAFELKGDIMFEGATEEEKLSRKMMKYATAPARTGNPNNDLSLMPAYNL 521
Db      468 DEAFVFGGAFELKGDIMFEGATEEEKLSRKMMKYATAPARTGNPNNDLSLMPAYNL 521
Qy      522 EGYLIQIGATTQQAOKLKEKEVAFWSE-----LLAKRLHA 556
Db      522 EGYLIQIGATTQQAOKLKEKEVAFWSE-----LLAKRLHA 556

RESULT 15
S10367
carboxylesterase (EC 3.1.1.1) ES-10 precursor, microsomal - rat
N:Alternate names: hydrolase A
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1991 #sequence _revision 30-Jun-1991 #text_change 09-Jul-2004
C:Accession: S10367; S12468; S51202; S23460; S14361
R:Robbi, M.; Beaufay, H.; Octave, J.N.
Biochem. J. 269, 451-458, 1990
A:Title: Nucleotide sequence of cDNA coding for rat liver pl 6.1 esterase (ES-10), a cart
A:Reference number: S10367; MUID:90351366; PMID:2286485
A:Accession: S10367
A:Molecule type: mRNA
A:Residues: 1-565 <ROB1>
A:Cross-references: UNIPROT:Q9R135; EMBL:X51974
A:Note: 168-Gln, 247-Tyr, 423-Met, and 506-Asn were also found
A:Note: the sequence from Fig. 4 is inconsistent with that from Fig. 5 in having 265-Tyr,
R:Robbi, M.
submitted to the EMBL Data Library, February 1990
A:Reference number: S12468
A:Accession: S12468
A:Molecule type: mRNA
A:Residues: 1-264, 'K', 266-565 <ROB2>
A:Cross-references: EMBL:X51974; NID:956898; PIDN:CAA36236.1; PID:956899
R:Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A:Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID:7986098
A:Accession: S51202
A:Molecule type: protein
A:Residues: 19-48 <MOR>
R:Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A:Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting ti
A:Reference number: S23460; MUID:92299008; PMID:1606962
A:Accession: S23460
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-185, 'Q', 187-422, 'W', 424-505, 'N', 507-565 <MED>
A:Cross-references: EMBL:X65296; NID:957553; PIDN:CAA46391.1; PID:957554
R:Gustad, R.; Sletten, K.; Lovhaug, D.; Fonnum, F.
Biochem. J. 274, 693-697, 1991
A:Title: Purification and characterization of carboxylesterases from rat lung.
A:Reference number: S14361; MUID:91190080; PMID:2012599
A:Accession: S14361
A:Molecule type: protein
A:Residues: 19-26, 'D', 28-37 <GAN>
C:Superfamily: Cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-565/Product: carboxylesterase #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CHES>
F:79,489/Binding site: carbonyl group (Asn) (covalent) #status predicted

```

F:221,466/active site: Ser, His #status predicted

Query Match 36.4%; Score 1121.5; DB 2; Length 565;

Best Local Similarity 42.4%; Pred. No. 3.6e-80; Mismatches 245; Conservative 79; Indels 55; Gaps 11;

```

QY 12 WCFPLIQLPLGHQMGKTGPSABGPQNTIRLGMIOGQVTVLGSPPVNVFLGVFPAP 71
DB 8 WLF-----LAACITMGY--PSSP-PVNTYVKGKVLGKYVNLGSPAQVAVFLGIFPAK 58
QY 72 PLGSLRFTNPQPASFWNDLREATSYENLCLQN-----SEWLLDQHMLKHYPKFGVS 124
DB 59 PLGSLRFPAPQPABFWNFVKNTTSYPMCSODAVGQVLSLFTNRKENIPLQF-----S 113
QY 125 EDCLYINITYAPAHADYTSKLPVLVWFPGAFKTSASIFDSALAAVEDLVVVVQYRLG 184
DB 114 EDCLYINITYPADLTKNSRLPVMWVHGGGLVVGASTYDQVLSAHENVVVTTIQYRLG 173
QY 185 IFGFPTTWDOHAPGNMAFKDQVAALSWYQKNIEFPGDPSVTTIFGESAGAISVSLLIS 244
DB 174 IWGFPTSTDEHSRGMGHLDQVAALHWQDNINANFGNPGSVTTIFGESAGGSVALVLS 233
QY 245 PMAGLFPKAIMESGVALLPYLEAHDYKESBDLQVVAHFCGNNASDSEALLRCLRTKPSK 304
DB 234 PLAKNLFHRAISESGVLTSLALITDSKRIANL--IATLSGCKTTTSAVMVYCLRQKTED 291
QY 305 ELTLISQKTSF-----TRVVDGAFFNEPDLDSQKAFKAIPIIIGVNN 349
DB 292 ELLETSLKLNLFKLDLGNPKESYPFLPTVIDGVVLPKTPBEILAEKSFNTVPYIVGINK 351
QY 350 HECGFLLPMKEAPILLSG-----SNKSLALHLIQNLIHPQYLHLVANEYFHDGHSLTE 404
DB 352 QEFGWIIPTLMGYPLSEKLDQKTAKS-LNKSYPYLKISEKMIPIVAEKIFGCTDDPAK 410
QY 405 IRDSILDLGDFVFPVVPALITARYHRDAGAPYFYEFRRHPQCFEDTKPAFYKADHADEV 464
DB 411 RKDLFQDLVAVLIVGVPVWVSRSHRDAGAPTFMYEFYRPSFVSAMRPKTVIGDHGDEL 470
QY 465 RFVFGGAFKGDIVFEGATEEBEKLRSKMKKYATTPARTGNPNGNDSLMPAYNLTEQY 524
DB 471 FSVFGSPFLK-----DSASEETNLSKVMKYMANFARNGSPNGGGLPHMPEYDQKEGY 524
QY 525 LQDLNMSIAGRLKEPRVDFTSTIPLIISASDMLHSP 562
DB 525 LKIGASTQAQRLKDKYVAFWSE-----LRAKKAABEP 557

```

Search completed: June 15, 2005, 09:39:49
Job time : 27 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: June 15, 2005, 09:39:26 ; Search time 86 Seconds
(without alignments)
2589.736 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079

Sequence: 1 MPOGLTSSASQWCFLLIQP.....PLSSLTFLSLQPPFFFCAP 581

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum March 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	581	US-10-023-515-2	Sequence 2, Appl1
2	3079	100.0	581	US-10-674-636-2	Sequence 2, Appl1
3	3079	100.0	581	US-10-757-262-46	Sequence 46, Appl1
4	2915	94.7	575	US-10-451-168-91	Sequence 91, Appl1
5	2900	94.2	642	US-10-433-256-10	Sequence 10, Appl1
6	2607.5	84.7	581	US-10-451-168-92	Sequence 2375, Ap
7	2606	84.6	525	US-10-094-749-2375	Sequence 196, Ap
8	2597.5	84.4	581	US-10-114-270-196	Sequence 2, Appl1
9	2588.5	84.1	618	US-10-381-898-2	Sequence 2219, Ap
10	2454	79.7	469	US-10-104-047-2219	Sequence 2, Appl1
11	1985	64.5	542	US-10-233-933A-2	Sequence 2, Appl1

12	1985	64.5	542	US-10-233-933A-4	Sequence 4, Appl1
13	1780	57.8	356	US-10-451-168-93	Sequence 99, Appl1
14	1230	39.9	549	US-09-925-298-689	Sequence 689, App
15	1230	39.9	549	US-10-102-806-689	Sequence 689, App
16	1230	39.9	550	US-10-267-756-5	Sequence 5, Appl1
17	1230	39.9	559	US-10-858-271-28	Sequence 28, Appl1
18	1186.5	38.5	554	US-09-895-860-4	Sequence 4, Appl1
19	1186.5	38.5	554	US-10-377-072-4	Sequence 4, Appl1
20	1186.5	38.5	554	US-10-377-072-4	Sequence 23, Appl1
21	1179.5	38.3	571	US-09-931-836-23	Sequence 1090, Ap
22	1179.5	38.3	571	US-10-036-342-23	Sequence 23, Appl1
23	1179.5	38.3	571	US-10-036-041-23	Sequence 23, Appl1
24	1179.5	38.3	571	US-10-028-072-542	Sequence 542, App
25	1179.5	38.3	571	US-10-028-072-542	Sequence 23, Appl1
26	1179.5	38.3	571	US-10-035-855-23	Sequence 542, App
27	1179.5	38.3	571	US-10-140-808-542	Sequence 542, App
28	1179.5	38.3	571	US-10-121-049-542	Sequence 542, App
29	1179.5	38.3	571	US-10-123-904-542	Sequence 542, App
30	1179.5	38.3	571	US-10-140-470-542	Sequence 542, App
31	1179.5	38.3	571	US-10-175-746-542	Sequence 542, App
32	1179.5	38.3	571	US-10-176-918-542	Sequence 542, App
33	1179.5	38.3	571	US-10-176-921-542	Sequence 542, App
34	1179.5	38.3	571	US-10-227-884-210	Sequence 210, App
35	1179.5	38.3	571	US-10-036-214-23	Sequence 23, Appl1
36	1179.5	38.3	571	US-10-137-865-542	Sequence 542, App
37	1179.5	38.3	571	US-10-140-474-542	Sequence 23, Appl1
38	1179.5	38.3	571	US-10-035-719-23	Sequence 542, App
39	1179.5	38.3	571	US-10-142-431-542	Sequence 542, App
40	1179.5	38.3	571	US-10-143-114-542	Sequence 210, App
41	1179.5	38.3	571	US-10-230-163-210	Sequence 210, App
42	1179.5	38.3	571	US-10-036-160-23	Sequence 23, Appl1
43	1179.5	38.3	571	US-10-230-338-210	Sequence 210, App
44	1179.5	38.3	571	US-10-142-419-542	Sequence 542, App
45	1179.5	38.3	571	US-10-218-631-210	Sequence 210, App

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Sltos-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Query Match 100.0%; Score 3079; DB 13; Length 581;
Best Local Similarity 100.0%; Pred. No. 1,1e-277;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPOGLTSSASQWCFLLIQP...PLSSLTFLSLQPPFFFCAP 60
DB 1 MPOGLTSSASQWCFLLIQP...PLSSLTFLSLQPPFFFCAP 60
QY 61 NVFLGVPAAPPLGSLRFTNPQAPSPWNLREATYSPIVLCIONSEMLLDQMLKVHYPK 120

```

Db      61 NVFLGVFAAPPLGSLRFTNPPQBPASPDNDLREATSYNCLQNSEWLLDQHLKHYPRK 120
Qy      121 FGVSBDCLYLNTIYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
Db      121 FGVSBDCLYLNTIYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
Qy      181 YRLGIFGFPTTMDQHPGNMFAFKQVVALSWQXNIEFGGDPSSVTIFGSSAGASIVSS 240
Db      181 YRLGIFGFPTTMDQHPGNMFAFKQVVALSWQXNIEFGGDPSSVTIFGSSAGASIVSS 240
Qy      241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDEKESDLOVVAHFCGNNASDBEALLRCURT 300
Db      241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDEKESDLOVVAHFCGNNASDBEALLRCURT 300
Qy      301 KPSKELLTLOSOKTSFTRVVDGAFPPNEBDLISQAKFAKIPSIIGVNNHCGFLPMKE 360
Db      301 KPSKELLTLOSOKTSFTRVVDGAFPPNEBDLISQAKFAKIPSIIGVNNHCGFLPMKE 360
Qy      361 APEILSGSNKSLAHLIQLNIHIPPQYLHVLVANEYFHDGSLTEIRDSLDDLDGDFPVV 420
Db      361 APEILSGSNKSLAHLIQLNIHIPPQYLHVLVANEYFHDGSLTEIRDSLDDLDGDFPVV 420
Qy      421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
Db      421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
Qy      481 EGATEEBEKLSSRKKMKYATFARTGNPNNDLSLMPAYNLTEOYLQDLNMSLGQRLKEP 540
Db      481 EGATEEBEKLSSRKKMKYATFARTGNPNNDLSLMPAYNLTEOYLQDLNMSLGQRLKEP 540
Qy      541 RVDFTWTSTIPLILSASDMLHSPSLSTFLSLQPFPPFCAP 581
Db      541 RVDFTWTSTIPLILSASDMLHSPSLSTFLSLQPFPPFCAP 581

```

```

RESULT 2
US-10-674-636-2
; Sequence 2, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Slioh-Santiago, Imaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT FILING DATE: US/10/674,636
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-674-636-2

```

```

Query Match      100.0%; Score 3079; DB 15; Length 581;
Best Local Similarity 100.0%; Pred. No. 1.le-277;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MPOGLTSSASQWCFLLIPLILGHROGKGPSAEGFORNRLGLWIGQKQVTVLGSVPV 60
Db      1 MPOGLTSSASQWCFLLIPLILGHROGKGPSAEGFORNRLGLWIGQKQVTVLGSVPV 60
Qy      61 NVFLGVFAAPPLGSLRFTNPPQBPASPDNDLREATSYNCLQNSEWLLDQHLKHYPRK 120
Db      61 NVFLGVFAAPPLGSLRFTNPPQBPASPDNDLREATSYNCLQNSEWLLDQHLKHYPRK 120

```

```

Qy      121 FGVSBDCLYLNTIYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
Db      121 FGVSBDCLYLNTIYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
Qy      181 YRLGIFGFPTTMDQHPGNMFAFKQVVALSWQXNIEFGGDPSSVTIFGSSAGASIVSS 240
Db      181 YRLGIFGFPTTMDQHPGNMFAFKQVVALSWQXNIEFGGDPSSVTIFGSSAGASIVSS 240
Qy      241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDEKESDLOVVAHFCGNNASDBEALLRCURT 300
Db      241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDEKESDLOVVAHFCGNNASDBEALLRCURT 300
Qy      301 KPSKELLTLOSOKTSFTRVVDGAFPPNEBDLISQAKFAKIPSIIGVNNHCGFLPMKE 360
Db      301 KPSKELLTLOSOKTSFTRVVDGAFPPNEBDLISQAKFAKIPSIIGVNNHCGFLPMKE 360
Qy      361 APEILSGSNKSLAHLIQLNIHIPPQYLHVLVANEYFHDGSLTEIRDSLDDLDGDFPVV 420
Db      361 APEILSGSNKSLAHLIQLNIHIPPQYLHVLVANEYFHDGSLTEIRDSLDDLDGDFPVV 420
Qy      421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
Db      421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
Qy      481 EGATEEBEKLSSRKKMKYATFARTGNPNNDLSLMPAYNLTEOYLQDLNMSLGQRLKEP 540
Db      481 EGATEEBEKLSSRKKMKYATFARTGNPNNDLSLMPAYNLTEOYLQDLNMSLGQRLKEP 540
Qy      541 RVDFTWTSTIPLILSASDMLHSPSLSTFLSLQPFPPFCAP 581
Db      541 RVDFTWTSTIPLILSASDMLHSPSLSTFLSLQPFPPFCAP 581

```

```

RESULT 3
US-10-757-262-46
; Sequence 46, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karichet, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 5164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; FILE REFERENCE: MP03-007PILNOMNIM
; CURRENT APPLICATION NUMBER: US/10/757,262
; PRIOR FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332

```


PRIOR FILING DATE: 2003-09-26
 NUMBER OF SEQ ID NOS: 136
 SOFTWARE: FASTSEQ for windows Version 4.0
 SEQ ID NO 46
 LENGTH: 581
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-757-262-46

Query Match 100.0%; Score 3079; DB 16; Length 581;
 Best Local Similarity 100.0%; Pred. No. 1,1e-277;
 Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPOGTTSSASQWCFPLITLQPLIGHKQMGKTPSABGPQNTRLGWIQKQVTVLSSPVV 60
 DB 1 MPOGTTSSASQWCFPLITLQPLIGHKQMGKTPSABGPQNTRLGWIQKQVTVLSSPVV 60
 QY 61 NVFLGVPFAAPPLGSLRFTNPQASPMNLRATSYPNLCQNSWMLLDQMLKVHYPK 120
 DB 61 NVFLGVPFAAPPLGSLRFTNPQASPMNLRATSYPNLCQNSWMLLDQMLKVHYPK 120
 QY 121 FGVSDECLYLYNYPADADTGSKLPLYVWFGGAKTGSASIFDSALAAVEDVLYVYVQ 180
 DB 121 FGVSDECLYLYNYPADADTGSKLPLYVWFGGAKTGSASIFDSALAAVEDVLYVYVQ 180
 QY 121 FGVSDECLYLYNYPADADTGSKLPLYVWFGGAKTGSASIFDSALAAVEDVLYVYVQ 180
 DB 121 FGVSDECLYLYNYPADADTGSKLPLYVWFGGAKTGSASIFDSALAAVEDVLYVYVQ 180
 QY 181 YRLGIFGFTTWDQHPGNMAFKDQVAAALSWQKNIIEFGDPSVTLFGESAGASIVSS 240
 DB 181 YRLGIFGFTTWDQHPGNMAFKDQVAAALSWQKNIIEFGDPSVTLFGESAGASIVSS 240
 QY 241 LILSPMAGLPHKAIMESGVALIPLYEADHYKSEDLQVVAHFCGNNASDSBALRCLRT 300
 DB 241 LILSPMAGLPHKAIMESGVALIPLYEADHYKSEDLQVVAHFCGNNASDSBALRCLRT 300
 QY 301 KPSKELLTLOKTKSFTRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHECGFLPMKE 360
 DB 301 KPSKELLTLOKTKSFTRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHECGFLPMKE 360
 QY 301 KPSKELLTLOKTKSFTRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHECGFLPMKE 360
 DB 301 KPSKELLTLOKTKSFTRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHECGFLPMKE 360
 QY 361 APELLSGNKSIALHLIQLNIHLIIPQYLLHVNEXFHDGSLTEIRDSLLDLGVPVAV 420
 DB 361 APELLSGNKSIALHLIQLNIHLIIPQYLLHVNEXFHDGSLTEIRDSLLDLGVPVAV 420
 QY 421 PALITARHARDAGAVFYEFRRHPCQCEDTKPAFYKADHADVRFVFGAFLKGDIVWF 480
 DB 421 PALITARHARDAGAVFYEFRRHPCQCEDTKPAFYKADHADVRFVFGAFLKGDIVWF 480
 QY 481 EGATEEEKLRSKMKKYATPARTGNPNDLSLMPAYNLTEQYQLDLNMSLGRLKEP 540
 DB 481 EGATEEEKLRSKMKKYATPARTGNPNDLSLMPAYNLTEQYQLDLNMSLGRLKEP 540
 QY 541 RVDFTSTTIPILSASDMLHSPSLTFLSLQPPFFCAP 581
 DB 541 RVDFTSTTIPILSASDMLHSPSLTFLSLQPPFFCAP 581

RESULT 4
 US-10-451-168-91
 Sequence 91, Application US/10451168
 Publication No. US20040091969A1
 GENERAL INFORMATION:
 APPLICANT: SMITHKLINE BEECHAM CORPORATION
 APPLICANT: SMITHKLINE BEECHAM P.L.C.
 APPLICANT: GLAXO GROUP LIMITED
 TITLE OF INVENTION: NOVEL COMPOUNDS
 FILE REFERENCE: GP50039
 CURRENT APPLICATION NUMBER: US/10/451,168
 PRIOR FILING DATE: 2003-11-12
 PRIOR APPLICATION NUMBER: PCT/US01/49232
 PRIOR FILING DATE: 2000-12-17
 PRIOR APPLICATION NUMBER: 60/256, 710
 PRIOR FILING DATE: 2000-12-19
 PRIOR APPLICATION NUMBER: 60/257, 048
 PRIOR FILING DATE: 2000-12-20
 PRIOR APPLICATION NUMBER: 60/260, 482

File 168

PRIOR FILING DATE: 2001-01-09
 PRIOR APPLICATION NUMBER: 60/264, 922
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: 60/266, 797
 PRIOR FILING DATE: 2001-02-06
 PRIOR APPLICATION NUMBER: 60/276, 988
 PRIOR FILING DATE: 2001-03-19
 PRIOR APPLICATION NUMBER: 60/281, 535
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/289, 622
 PRIOR FILING DATE: 2002-06-28
 NUMBER OF SEQ ID NOS: 110
 SOFTWARE: FASTSEQ for windows Version 4.0
 SEQ ID NO 91
 LENGTH: 575
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-451-168-91

Query Match 94.7%; Score 2915; DB 15; Length 575;
 Best Local Similarity 97.0%; Pred. No. 2,1e-262;
 Matches 553; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 12 WCFPLITLQPLIGHKQMGKTPSABGPQNTRLGWIQKQVTVLSSPVVNVFLGVPFAAP 71
 DB 14 WAIWLAAP-----TKGPSABGPQNTRLGWIQKQVTVLSSPVVNVFLGVPFAAP 65
 QY 72 PLGSLRFTNPQASPMNLRATSYPNLCQNSWMLLDQMLKVHYKFGVSEDECLYLN 131
 DB 66 PLGSLRFTNPQASPMNLRATSYPNLCQNSWMLLDQMLKVHYKFGVSEDECLYLN 125
 QY 132 IYAPAHADTGSKLPLYVWFGGAKTGSASIFDSALAAVEDVLYVYVQYRLGIFGFTT 191
 DB 126 IYAPAHADTGSKLPLYVWFGGAKTGSASIFDSALAAVEDVLYVYVQYRLGIFGFTT 185
 QY 192 WDOHAPGNMAFKDQVAAALSWQKNIIEFGDPSVTLFGESAGASIVSSLILSPMAGL 251
 DB 186 WDOHAPGNMAFKDQVAAALSWQKNIIEFGDPSVTLFGESAGASIVSSLILSPMAGL 245
 QY 252 HKAIMESGVALIPLYEADHYKSEDLQVVAHFCGNNASDSBALRCLRTKPSKELLTLO 311
 DB 246 HKAIMESGVALIPLYEADHYKSEDLQVVAHFCGNNASDSBALRCLRTKPSKELLTLO 305
 QY 312 KTKSFTRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHECGFLPMKEAPEILSGNKS 371
 DB 306 KTKSFTRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHECGFLPMKEAPEILSGNKS 365
 QY 372 LALHLIQLNIHLIIPQYLLHVNEXFHDGSLTEIRDSLLDLGDFVFPALITARYARD 431
 DB 366 LALHLIQLNIHLIIPQYLLHVNEXFHDGSLTEIRDSLLDLGDFVFPALITARYARD 425
 QY 432 AGAPVYFYEFRRHPCQCEDTKPAFYKADHADVRFVFGAFLKGDIVWFEGETEEKLS 491
 DB 426 AGAPVYFYEFRRHPCQCEDTKPAFYKADHADVRFVFGAFLKGDIVWFEGETEEKLS 485
 QY 492 RRMKMYATPARTGNPNDLSLMPAYNLTEQYQLDLNMSLGRLKEPRVDFTSTTIP 551
 DB 486 RRMKMYATPARTGNPNDLSLMPAYNLTEQYQLDLNMSLGRLKEPRVDFTSTTIP 545
 QY 552 ILSASDMLHSPSLTFLSLQPPFFCAP 581
 DB 546 ILSASDMLHSPSLTFLSLQPPFFCAP 575

RESULT 5
 US-10-433-256-10
 Sequence 10, Application US/10433256
 Publication No. US20040081980A1
 GENERAL INFORMATION:
 APPLICANT: SANUCAMALIA, Madhusudan M.; YAO, Monique G.
 APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
 APPLICANT: ARVIZU, Chandra S.; RING, Huijun Z.
 APPLICANT: LEE, Ernestine A.; DING, Li

X

```

/ APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
/ APPLICANT: YUE, Henry; TRIBOULEY, Catherine M.
/ APPLICANT: LU, Dying Aina M.; LAU, Preeti G.
/ APPLICANT: WARREN, Bridget A.; YANG, Junning
/ APPLICANT: CHAMLA, Narinder K.; NGUYEN, Daniel B.
/ APPLICANT: GANDHI, Ameena R.; LU, Yan
/ APPLICANT: ISON, Craig H.
/ TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
/ FILE REFERENCE: PI-0113 USN
/ CURRENT APPLICATION NUMBER: US/10/433,256
/ PRIOR FILING DATE: 2003-05-30
/ PRIOR APPLICATION NUMBER: PCT/US01/47429
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: US 60/254,308
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: US 60/256,189
/ PRIOR FILING DATE: 2000-12-15
/ PRIOR APPLICATION NUMBER: US 60/257,713
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: US 60/262,706
/ PRIOR FILING DATE: 2001-01-19
/ PRIOR APPLICATION NUMBER: US 60/266,020
/ PRIOR FILING DATE: 2001-02-02
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PERL Program
/ SEQ ID NO 10
/ LENGTH: 642
/ TYPE: PRF
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No: 6538080CD1
US-10-433-256-10

```

211/29/00

```

Query Match      94.2%; Score 2900; DB 15; Length 642;
Best Local Similarity 99.5%; Pred. No. 6,3e-261;
Matches 548; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 GPSAGPQPNTRLGMIOGKQVTVLGSPPVNVFLGPPAAPPLGSLRFTNPQSPMNDL
DB 92 GPSAGPQPNTRLGMIOGKQVTVLGSPPVNVFLGPPAAPPLGSLRFTNPQSPMNDL 151
QY 91 REATSPNLCIONSEWMLLDQHLKVHYPKGVSEDCIXNTIYAPAHADTSSKLPVYWF
DB 152 REATSPNLCIONSEWMLLDQHLKVHYPKGVSEDCIXNTIYAPAHADTSSKLPVYWF 211
QY 151 PGGAFTKTSASIFDGSALAAYEDVLVVVQYRLGIFGFTTMDQHPGNMAFKQVVALS
DB 212 PGGAFTKTSASIFDGSALAAYEDVLVVVQYRLGIFGFTTMDQHPGNMAFKQVVALS 271
QY 211 MVQKNIEFFGDDPSVITFGESAGAISSVSLILSPMAKGLFHKAIMESGVALIPYLEAND
DB 272 MVQKNIEFFGDDPSVITFGESAGAISSVSLILSPMAKGLFHKAIMESGVALIPYLEAND 331
QY 271 YEKSEDLQVVAHFPGNNASDSEALLRCRTKPSKELTLISQKTSFTRVVDGAFEPNEPL
DB 332 YEKSEDLQVVAHFPGNNASDSEALLRCRTKPSKELTLISQKTSFTRVVDGAFEPNEPL 391
QY 331 DLSQKAFKAIPIISIIIGVNNHCGFLIPMKEAPELILSSNKSALHLIIONILHIPPQYLHL
DB 392 DLSQKAFKAIPIISIIIGVNNHCGFLIPMKEAPELILSSNKSALHLIIONILHIPPQYLHL 451
QY 391 VANEYFHDKSLTEIRDSLLDLGDFVFPVVPALITAYHRDAGAPVYFFRRRPOCFED
DB 452 VANEYFHDKSLTEIRDSLLDLGDFVFPVVPALITAYHRDAGAPVYFFRRRPOCFED 511
QY 451 TKPAFYVADHAEVRFVFGAFLKGDIVMEGATEEBEKLISRKMKYKWAFFATGPNNGN
DB 512 TKPAFYVADHAEVRFVFGAFLKGDIVMEGATEEBEKLISRKMKYKWAFFATGPNNGN 571
QY 511 DLSLMPVYNLTQYLOLDLMSIGORLKEPRVDFWSTTIPILISASDMLHSPILSITFLS
DB 572 DLSLMPVYNLTQYLOLDLMSIGORLKEPRVDFWSTTIPILISASDMLHSPILSITFLS 631

```

```

QY 571 LLOPFFPCAP 581
DB 632 LLOPFFPCAP 642

```

```

RESULT 6
US-10-451-168-92
/ Sequence 92, Application US/10451168
/ Publication No. US20040091969A1
/ GENERAL INFORMATION:
/ APPLICANT: SMITHKLINE BEECHAM CORPORATION
/ APPLICANT: GLAXO GROUP LIMITED
/ APPLICANT: SMITHKLINE BEECHAM P.L.C.
/ TITLE OF INVENTION: NOVEL COMPOUNDS
/ FILE REFERENCE: GP50039
/ CURRENT APPLICATION NUMBER: US/10/451,168
/ PRIOR FILING DATE: 2003-11-12
/ PRIOR APPLICATION NUMBER: PCT/US01/49232
/ PRIOR FILING DATE: 2000-12-17
/ PRIOR APPLICATION NUMBER: 60/256,710
/ PRIOR FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 60/257,048
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 60/260,482
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: 60/264,922
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/266,797
/ PRIOR FILING DATE: 2001-02-06
/ PRIOR APPLICATION NUMBER: 60/276,988
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/281,535
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/289,622
/ PRIOR FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 92
/ LENGTH: 581
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-451-168-92

```

```

Query Match      84.7%; Score 2607.5; DB 15; Length 581;
Best Local Similarity 91.4%; Pred. No. 1e-233;
Matches 499; Conservative 7; Mismatches 25; Indels 15; Gaps 3;

QY 12 WCFELLQPLGHRQWKTGPSAGPQPNTRLGMIOGKQVTVLGSPPVNVFLGPPAAP
DB 14 WAIWVAAP-----TKGPSAGPQPNTRLGMIOGKQVTVLGSPPVNVFLGPPAAP 65
QY 72 PLGSRLFTNPQSPMNDLREATSPNLCIONSEWMLLDQHLKVHYPKGVSEDCIXNTI
DB 66 PLGSRLFTNPQSPMNDLREATSPNLCIONSEWMLLDQHLKVHYPKGVSEDCIXNTI 125
QY 132 IYAPAHADTSSKLPVYWFPGGAFTKTSASIFDGSALAAYEDVLVVVQYRLGIFGFTT
DB 126 IYAPAHADTSSKLPVYWFPGGAFTKTSASIFDGSALAAYEDVLVVVQYRLGIFGFTT 185
QY 192 WQOHAPEGNAFQVVAALSWQKNIEFFGDDPSVITFGESAGAISSVSLILSPMAKGLF
DB 186 WQOHAPEGNAFQVVAALSWQKNIEFFGDDPSVITFGESAGAISSVSLILSPMAKGLF 251
QY 251 HKAIMESGVALIPYLEANDYEKSEDLQVVAHFPGNNASDSEALLRCRTKPSKELTLISQ
DB 246 HKAIMESGVALIPYLEANDYEKSEDLQVVAHFPGNNASDSEALLRCRTKPSKELTLISQ 305
QY 312 KTKSFTRVVDGAFEPNEPLDLSQKAFKAIPIISIIIGVNNHCGFLIPMKE--APEILSGSN
DB 306 KTKSFTRVVDGAFEPNEPLDLSQKAFKAIPIISIIIGVNNHCGFLIPMKE--APEILSGSN 369
QY 369 KTKSFTRVVDGAFEPNEPLDLSQKAFKAIPIISIIIGVNNHCGFLIPMKE--APEILSGSN
DB 370 KTKSFTRVVDGAFEPNEPLDLSQKAFKAIPIISIIIGVNNHCGFLIPMKE--APEILSGSN 424

```

```

Db      : 366 RDALASTAGHFRHQRHLPQYLHVANVEYHDKSLTEIRDSLLDLDGVFFVVPALII 425
Qy      : 425 TARYHRDGAIVYFVEFRHRPOCFEDTTPAPYKADHAEVRFVFGALDKGDIVNFEAGT 484
Db      : 426 TARYHRDGAIVYFVEFRHRPOCFEDTTPAPYKADHAEVRFVFGALDKGDIVNFEAGT 485
Qy      : 485 EEEKLSKMKMYATFARTGNPNNGNDLSMPAYNLTEQYQLDLMNSLGORLKEPRVD 544
Db      : 486 EEEKLSKMKMYATFARTGNPNNGNDLSMPAYNLTEQYQLDLMNSLGORLKEPRVD 545
Qy      : 545 WTSTIP 550
Db      : 546 WVTGYP 551

```

RESULT 7

```

US-10-094-749-2375
; Sequence 2375, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAMI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAORIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 08435/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2375
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2375

```

```

Query Match      84.6%; Score 2606; DB 15; Length 525;
Best Local Similarity 90.9%; Pred. No. 1,2e-233;
Matches 501; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

```

```

Qy      : 31 GPSAEGPQNRRLGWIQGGQVTVLGSPPVNVNVLGVPAAAPLGLRFTNPPQASPMWNL 90
Db      : 25 GPSAEGPQNRRLGWIQGGQVTVLGSPPVNVNVLGVPAAAPLGLRFTNPPQASPMWNL 84
Qy      : 91 REATSYPMCLQNSWMLLDQMLKVHYPKFGVSDCLYLNLYADAHADTGSKLFLVWVF 150
Db      : 85 REATSYPMCLQNSWMLLDQMLKVHYPKFGVSDCLYLNLYADAHADTGSKLFLVWVF 144
Qy      : 151 PGGAFTKTSASIFDSALAAVEDVTVVVVQVRLGIRGFTTMDQAPGMAFKDQVALIS 210
Db      : 145 PGGAFTKTSASIFDSALAAVEDVTVVVVQVRLGIRGFTTMDQAPGMAFKDQVALIS 204
Qy      : 211 WVKNIREFGGDPSVTIFGSAGAISSVSLISPMAGLFRKAIMESGVAIIPYLEAHD 270

```

```

Db      : 205 WVKNIREFGGDPSVTIFGSAGAISSVSLISPMAGLFRKAIMESGVAIIPYLEAHD 264
Qy      : 271 YEKSEDLQVNAHFCGNNSDSEALLRCRTKPSKELLTSQTKSFTRVVDGAPFPNEPL 330
Db      : 265 YEKSEDLQVNAHFCGNNSDSEALLRCRTKPSKELLTSQTKSFTRVVDGAPFPNEPL 324
Qy      : 331 DLLSQAKFALPSIIIGVNHCEGFLPMKEAPEILSGNSKSLALHLIONIILHIPPQYLHL 390
Db      : 325 DLLSQAKFALPSIIIGVNHCEGFLPMKEAPEILSGNSKSLALHLIONIILHIPPQYLHL 384
Qy      : 391 VANEYFHDKSLTEIRDSLLDLDGVFFVVPALITARYHRDAGAVYFVEFRHRQCFED 450
Db      : 385 VANEYFHDKSLTEIRDSLLDLDGVFFVVPALITARYHRDAGAVYFVEFRHRQCFED 424
Qy      : 451 TKPAFVKADHAEVRFVFGAFLLKGDIVMFGATSEELLSRKMKMYATFARTGNPNGN 510
Db      : 425 -----BGATSEELLSRKMKMYATFARTGNPNGN 454
Qy      : 511 DLSLMPAYNLTEQYQLDLMNSLGORLKEPRVDFTSTIPILSADMHSPLSLTFLS 570
Db      : 455 DLSLMPAYNLTEQYQLDLMNSLGORLKEPRVDFTSTIPILSADMHSPLSLTFLS 514
Qy      : 571 LLQPPFFCAP 581
Db      : 515 LLQPPFFCAP 525

```

RESULT 8

```

US-10-114-270-196
; Sequence 196, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Virel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Paturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zernusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Burgess, Catherine B.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liere, Mario W.
; APPLICANT: Raesceilli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark B.
; TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906

```

```

; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining prior application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 196
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-196
```

```

Query Match      84.4%; Score 2597.5; DB 15; Length 581;
Best Local Similarity 94.3%; Pred. No. 8.7e-233;
Matches 497; Conservative 4; Mismatches 17; Indels 9; Gaps 3;
```

```

QY 31 GPSABGPQNTRLGWIQKQVTVLGSPPVNVFLGVPAAPLGSIRFTNPQASPMNDL 90
DB 27 GPSABGPQNTRLGWIQKQVTVLGSPPVNVFLGVPAAPLGSIRFTNPQASPMNDL 86
QY 91 REATSYPNLCIONSEWLLIDQHLKVHYPKFVSEDCLYLNTYAPAHADTSGKLPVLWF 150
DB 87 REATSYPNLCIONSEWLLIDQHLKVHYPKFVSEDCLYLNTYAPAHADTSGKLPVLWF 146
QY 151 PGAFKTSASIFDGSALAAVEDVLVWVQYRLGIFGFTTWDQHPGMAFKDOVALS 210
DB 147 PGAFKTSASIFDGSALAAVEDVLVWVQYRLGIFGFTTWDQHPGMAFKDOVALS 206
QY 211 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAH 270
DB 207 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAH 266
QY 271 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKELTLTSSQKTSFTRVNDGAPFPNEPL 330
DB 267 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKELTLTSSQKTSFTRVNDGAPFPNEPL 326
QY 331 DLSQKAFKAIPSIIGVNNHECGFLPMKE--APEILSGSNKSLAL-----HLIQNLIHI 383
DB 327 DLSQKAFKAIPSIIGVNNHECGFLPMKE--APEILSGSNKSLAL-----HI 384
QY 384 PPOYLHVANEYFHDKSLTEIRDSLLDLGDVFFVPVALLTARYHRDAGAPVYFEFRH 443
DB 385 PPOYLHVANEYFHDKSLTEIRDSLLDLGDVFFVPVALLTARYHRDAGAPVYFEFRH 444
QY 444 RPOCFEDTKPAFVADHADEVRFVGGAFKGLKDIYMEGATEEBEKLSSRKMKWTAFAR 503
DB 445 RPOCFEDTKPAFVADHADEVRFVGGAFKGLKDIYMEGATEEBEKLSSRKMKWTAFAR 504
QY 504 TGNNGNDLSIMPAYNLTBOYLQDLDMNSLGQRLKEPRRDVWVWGYTP 550
DB 505 TGNNGNDLSIMPAYNLTBOYLQDLDMNSLGQRLKEPRRDVWVWGYTP 551
```

RESULT 9

```

US-10-381-898-2
; Sequence 2, Application US/10381898
; Publication No. US20040086887A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKI, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; IAL, Preeti G.;
```

```

; APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANMALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOLEV, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Koderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junning; YAO, Monique; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/236,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473645CD1
US-10-381-898-2
```

```

Query Match      84.1%; Score 2588.5; DB 15; Length 618;
Best Local Similarity 94.4%; Pred. No. 6.6e-232;
Matches 491; Conservative 1; Mismatches 5; Indels 23; Gaps 1;
```

```

QY 31 GPSABGPQNTRLGWIQKQVTVLGSPPVNVFLGVPAAPLGSIRFTNPQASPMNDL 90
DB 92 GPSABGPQNTRLGWIQKQVTVLGSPPVNVFLGVPAAPLGSIRFTNPQASPMNDL 151
QY 91 REATSYPNLCIONSEWLLIDQHLKVHYPKFVSEDCLYLNTYAPAHADTSGKLPVLWF 150
DB 152 REATSYPNLCIONSEWLLIDQHLKVHYPKFVSEDCLYLNTYAPAHADTSGKLPVLWF 211
QY 151 PGAFKTSASIFDGSALAAVEDVLVWVQYRLGIFGFTTWDQHPGMAFKDOVALS 210
DB 212 PGAFKTSASIFDGSALAAVEDVLVWVQYRLGIFGFTTWDQHPGMAFKDOVALS 271
QY 211 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAH 270
DB 272 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAH 331
QY 271 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKELTLTSSQKTSFTRVNDGAPFPNEPL 330
DB 332 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKELTLTSSQKTSFTRVNDGAPFPNEPL 391
QY 331 DLSQKAFKAIPSIIGVNNHECGFLPMKEAPEILSGSNKSLALHLIQNLIHIPOYLHL 390
DB 392 DLSQKAFKAIPSIIGVNNHECGFLPM-----HIPOYLHL 428
QY 391 VANEYFHDKSLTEIRDSLLDLGDVFFVPVALLTARYHRDAGAPVYFEFRHRPOCFED 450
DB 429 VANEYFHDKSLTEIRDSLLDLGDVFFVPVALLTARYHRDAGAPVYFEFRHRPOCFED 488
QY 451 TKPAVKAADHADEVRFVGGAFKGLKDIYMEGATEEBEKLSSRKMKWTAFARGNPGN 510
```

```
Db      489  TKPAFAVKADHADEVAFVFGAFLKGDIVMFEAGATEEKKLSRKMKMYATFARTGNPNGN 548
Qy      511  DLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFTSTIP 550
Db      549  DLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFTSTIP 588

RESULT 10
US-10-104-047-2219
; Sequence 2219, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2219
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2219

Query Match      79.7%; Score 2454; DB 15; Length 469;
Best Local Similarity 99.6%; Pred. No. 1.5e-219;
Matches 467; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      113  MLKHYHPKGVSEDCLYNTIYAPAHADTGSKLPLYVMPFGAFKTSASIFDGSALAYE 172
Db      1    MLKHYHPKGVSEDCLYNTIYAPAHADTGSKLPLYVMPFGAFKTSASIFDGSALAYE 60
Qy      173  DVLVVVVVYRIGIFGFTTWDQAHAPGNNAFKDQVAAALSMVQKNIEFGGDPSSVTIFGSS 232
Db      61  DVLVVVVVYRIGIFGFTTWDQAHAPGNNAFKDQVAAALSMVQKNIEFGGDPSSVTIFGSS 120
Qy      233  AGAISVSSLLISPMKAGLPHKAIMESGVAIIPYLEAHDEKSESDIQVVAHFCGNNAOSE 292
Db      121  AGAISVSSLLISPMKAGLPHKAIMESGVAIIPYLEAHDEKSESDIQVVAHFCGNNAOSE 180
Qy      293  ALALCLRTKPSKELLTTSOKTKSPTRVVDGAFPEPNEPLDLSOKFAIPSTIGVNNHRC 352
Db      181  ALALCLRTKPSKELLTTSOKTKSPTRVVDGAFPEPNEPLDLSOKFAIPSTIGVNNHRC 240
Qy      353  GFLPMPKAPRLLSGSNKSLALHLIQLNLIHTPPQYLHLVANEYFHDKSLTEIRDLLDL 412
Db      241  GFLPMPKAPRLLSGSNKSLALHLIQLNLIHTPPQYLHLVANEYFHDKSLTEIRDLLDL 300
Qy      413  LGDVEFVVVALITARYHRDAGAPVYFYEPRHRPQCFEDTKPAFAVKADHADEVAFVFGANF 472
Db      301  LGDVEFVVVALITARYHRDAGAPVYFYEPRHRPQCFEDTKPAFAVKADHADEVAFVFGANF 360
Qy      473  LKGDIVMEEGATEEKKLSRKMKMYATFARTGNNGNDLSLMPAYNLTEOYLQDLNMS 532
Db      361  LKGDIVMEEGATEEKKLSRKMKMYATFARTGNNGNDLSLMPAYNLTEOYLQDLNMS 420
Qy      533  LGORLKEPRVDFTSTIPILISASDMLHSPSLSTFLSLQFPFFFCAP 581
Db      421  LGORLKEPRVDFTSTIPILISASDMLHSPSLSTFLSLQFPFFFCAP 469

RESULT 11
US-10-233-933A-2
; Sequence 2, Application US/10233933A
; Publication No. US20040214171A1
; GENERAL INFORMATION:
; APPLICANT: Yamashita, Tetsuro
; APPLICANT: Miyazaki, Masao
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG FP02US006
```

```
; CURRENT APPLICATION NUMBER: US/10/233,933A
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: JP2002-057908
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Felis catus
US-10-233-933A-2

Query Match      64.5%; Score 1985; DB 16; Length 542;
Best Local Similarity 71.7%; Pred. No. 1e-175;
Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;

Qy      31  GPSAEGPQNRTRLGWITQSKQVTVLGSPPVNVVFLGVPPAAPPLGSLRTRTNPQASPMWNL 90
Db      22  GPADAPVRSSTRLGWVRGKQTVTVLGSPPVNVVFLGVPPAAPPLGSLRTRTNPQASPMWNL 81
Qy      91  REATSVPLCLQNSFWLLDDHMLKVHPKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 150
Db      82  RNATSVPLCLQNSFWLLDDHMLKVHPKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 141
Qy      151  PGAFKTSASIFDGSALAAVEDVLLVVVYRIGIFGFTTWDQAHAPGNNAFKDQVAAALS 210
Db      142  PGAFKTSASIFDGSALAAVEDVLLVVVYRIGIFGFTTWDQAHAPGNNAFKDQVAAALS 201
Qy      211  WVQKNIEFGGDPSSVTIFGSSAGAISSVSLISPMKAGLPHKAIMESGVAIIPYL-EAH 269
Db      202  WVRDNIIEFGGDPSSVTIFGSSAGAISSVSLISPMKAGLPHKAIMESGVAIIPYLMPRP 261
Qy      270  DYKSESDIQVVAHFCGNNAOSEALRLCLRTKPSKELLTTSOKTKSPTRVVDGAFPEPNEP 329
Db      262  GDERKDDIQVLAIRICGCHASDPAALLOCLTRAKPSELDISKLTSPIDIDFFPDEP 321
Qy      330  LDLSQKAFKAIPISTIGVNNHCGFLPMPKAPRLLSGSNKSLALHLIQLNLIHTPPQYLH 389
Db      322  VALLTQKAFNSVPSIIGVNNHCGFLPMPKAPRLLSGSNKSLALHLIQLNLIHTPPQYLH 380
Qy      390  LVANEYFHDKSLTEIRDLLDLGDFVFPVVALITARYHRDAGAPVYFYEPRHRPQCFE 449
Db      381  LVADHYFYNKSPVEIRDSPFLDLGDFVFPVVALITARYHRDAGAPVYFYEPRHRPQCFE 440
Qy      450  DTKPAFAVKADHADEVAFVFGAFLKGDIVMFEAGATEEKKLSRKMKMYATFARTGNPNGN 509
Db      441  DTKPAFAVKADHADEVAFVFGAFLKGDIVMFEAGATEEKKLSRKMKMYATFARTGNPNGN 500
Qy      510  NDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFTSTI 549
Db      501  EGVPLMPAYNLTEOYLQDLNMSLQRLKEPRVDFTSTI 540

RESULT 12
US-10-233-933A-4
; Sequence 4, Application US/10233933A
; Publication No. US20040214171A1
; GENERAL INFORMATION:
; APPLICANT: Yamashita, Tetsuro
; APPLICANT: Miyazaki, Masao
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG FP02US006
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: JP2002-057908
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Felis catus
US-10-233-933A-4
```

```

Query Match      64.5%; Score 1985; DB 16; Length 542;
Best Local Similarity 71.7%; Pred. No. 1e-175;
Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;

QY 31 GSAAGPQNRTRIGMIGKQVTVLGGSPVNVNVLGVFPFAAPPLGSLRFTNPQSPASPDNL 90
DB 22 GPAAADAPVSTRIGWRGKQTVLGGSTVPVNMFLGIPYAAPPLGRLRFPQPKALPGNDF 61
QY 91 REATSPNCLONSEWLLDDQHLKXHPKFGVSEDCYLYNTIYAPHAATGSLPYLVWF 150
DB 82 RNATSPKLCFQDLEWLVSYQHLKXRYPLENSEDCYLYNTIYAPHAADNGSLPYVWF 141
QY 151 PGCAFPTGSASIFDGSALAAVEDLVVVVQYRLGIFGFTTMDQHPGMAFDDQVAALS 210
DB 142 PGCAFPTGSASIFDGSALAAVEDLVVVQYRLGIFGFTTMDQHPGMAFDDQVAALT 201
QY 211 WYQKNIFFGDPSSVTITGESAGATSVSLLISPMAKGLFKHAIMESGVAIIPYL-EAH 269
DB 202 WVRNDIIEFGDPRSVTITGESAGATSVSLLISPIANGLFKHAIMESGVAIIPYLMRPP 261
QY 270 DYKSEDDLOVNAHFCGNNASDSEALLRCURTKEPSKELLTSQKTSFTRVVDGAFPPNRP 329
DB 262 GBERKDDLOVNAHFCGNNASDSEALLRCURTKEPSKELLTSQKTSFTRVVDGAFPPNRP 321
QY 330 LDLLSQKAFKALPSIIGVNNHCEGFLPMKEAPEILSGSNKSLALHLIQLNIHIPQYLH 389
DB 322 VALLTQKAFNSVPSIIGVNNHCEGFL-STESEILSGSNRSLALYVHTFNIPTQYLH 380
QY 390 LVANEFYHDKSLTEIRDSLLDLGDVFFVVPALITARYHRDAGAPVYFEFRHRPQCFE 449
DB 361 LVADHYFFNKHSPVIRDSFLDLGDVFFVVPALITARYHRDAGAPVYFEFRHRPQCLN 440
QY 450 DTKPAFVADHAEVRFVFGAFLKGDIVMEFEGATEEBSKLSRKMMKWATPAPRTGNNG 509
DB 441 DTRPAFVADHAEVRFVFGAFLKGDIVMEFEGATEEBSKLSRKMMKWATPAPRTGNNG 500
QY 510 NDLSLWPAVNTTEQYLQDLNNSLQGLRKEPRVDFWTSTI 549
DB 501 EGVPLMPAYTQSEQYTKIDLVSQKLEQDEVEFMNNTI 540

RESULT 13
US-10-451-168-93
; Sequence 93, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; PRIOR FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-93

Query Match      57.8%; Score 1780; DB 15; Length 356;
Best Local Similarity 100.0%; Pred. No. 7.1e-157;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ILSPMAKGLFKHAIMESGVAIIPYLEAHDEKESDLOVNAHFCGNNASDSEALLRCURT 301
DB 17 ILSPMAKGLFKHAIMESGVAIIPYLEAHDEKESDLOVNAHFCGNNASDSEALLRCURT 76
QY 302 PSKELLTSQKTSFTRVVDGAFPPNEPLDLSQKAFKALPSIIGVNNHCEGFLPMKEA 361
DB 77 PSKELLTSQKTSFTRVVDGAFPPNEPLDLSQKAFKALPSIIGVNNHCEGFLPMKEA 136
QY 362 PEILSGSNKSLALHLIQLNIHIPQYLHLVANEYFHDKSLTEIRDSLLDLGDVFFVVP 421
DB 137 PEILSGSNKSLALHLIQLNIHIPQYLHLVANEYFHDKSLTEIRDSLLDLGDVFFVVP 196
QY 422 ALITARYHRDAGAPVYFEFRHRPQCFEDTKPAFVKADHAEVRFVFGAFLKGDIVMEF 481
DB 197 ALITARYHRDAGAPVYFEFRHRPQCFEDTKPAFVKADHAEVRFVFGAFLKGDIVMEF 256
QY 482 GATEEBSKLSRKMMKWATPAPRTGNPNNDLSLMPAYNTTEQYLQDLNNSLQGLRKEPR 541
DB 257 GATEEBSKLSRKMMKWATPAPRTGNPNNDLSLMPAYNTTEQYLQDLNNSLQGLRKEPR 316
QY 542 VDFWTSTIIPILSASDMLSPSLSTLTLQPPFFFCAP 561
DB 317 VDFWTSTIIPILSASDMLSPSLSTLTLQPPFFFCAP 356

RESULT 14
US-09-925-298-689
; Sequence 689, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 689
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-689

Query Match      39.9%; Score 1230; DB 9; Length 549;
Best Local Similarity 46.7%; Pred. No. 2.6e-105;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

QY 31 GSAAGPQNRTRIGMIGKQVTVLGGSPVNVNVLGVFPFAAPPLGSLRFTNPQSPASPDNL 90
DB 16 GDSASPIRTHTGVLGSLVHVKGANAGVQTFGLIPPAKPEPLGLRAPPEPSPSGV 75

```

```

QY 91 REATSVPLCLQN-----SEWILLDOHMLKVHPKGVSEDCLYINIVAPAHADGSKLP 145
DB 76 RDGTHPAMCLODLTVAVESEFL-----SQFNMTEPDSMSBEDCLYISTYTPAHSHGSLNP 131
QY 146 VLVWFPFGAFKFGASIFDGSALAAVEDVLYVVOYRIGIFGFFTTMOAHAGNAFQDQ 205
DB 132 VVWVTHGGALVFGMASTLYDGSMTLALENVVVIYIYRLGVLDGFSTGDKHATGNMGYLDQ 191
QY 206 VVALSMVQKNIEFFGADPSVYTFGESAGATSVSSLISLSPMAKGLFHKAIMSGVAILIPY 265
DB 192 VVALRMVQONIAHFGNDRTYTFGESAGTSVSSLVSPISQGLFHGAIMSGVALLPBG 251
QY 266 LEADHYKSEDL--OVVAHFCGNNASDSEBALRCRTPKSKELLTLOKTKSFTRVVDGA 323
DB 252 LIA-----SSADVI STVVA NLSACDQVDSBALVGCIRGSKKEIILAINKPFKMI PCVVDGV 307
QY 324 FFPNEPDLISQAKAPKAPISITIGVNNHCGFLP-----MKEAPELISGSKSLAHLILQN 379
DB 308 FLPRHPOBILASADQPVPSIVGVNNNEFGWLI PKVMRIYDTQKEMDRASQALLOKMLT 367
QY 380 ILHIPPOYLILVANEYFHDKISLTERISDLIDLGDVFFVVPALITARYHRDAGAPVYFY 439
DB 368 LIMLPPTGDLREYIGDNGDPOTLOAQOFQEMMADSMFVLPALQVAHF-QCSRAPVYFY 426
QY 440 EFRHRPOCFEDTKPAFVADHADEVRFVFGAFLKGDIVMEGATEEBSKLSRKMMKTYA 499
DB 427 EFQHPSPWLKNI RPPHMKADHGDDEL PVPF-RSFCGNYIKF---TEEBEQLSRKMMKTYA 482
QY 500 TPARTEGNNGNDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDWTSTIP 550
DB 483 NFARNGNPNGBGLPMPPLFDQEBEYLOLNLQPAVGRALKAHRLQFWKXALP 533

```

RESULT 15
US-10-102-806-689

```

; Sequence 689, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PICI
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 689
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-689

```

Query Match 39.9%; Score 1230; DB 14; Length 549;
Best Local Similarity 46.7%; Pred. No. 2,6e-105;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

```

QY 31 GPSAEGPQNRTRLGNIGKQVTVLGSPPVNVNVLGVPFPAAPLGSIRFTNPPQASPMDNL 90
DB 16 GQDSASPIRTHTGTGVLGSLVHVKGANAGVQTFGLGIPFAKPPPLGPRFAPPEPSPWSGV 75
QY 91 REATSVPLCLQN-----SEWILLDOHMLKVHPKGVSEDCLYINIVAPAHADGSKLP 145

```

```

DB 76 RDGTHPAMCLODLTVAVESEFL-----SQFNMTEPDSMSBEDCLYISTYTPAHSHGSLNP 131
QY 146 VLVWFPFGAFKFGASIFDGSALAAVEDVLYVVOYRIGIFGFFTTMOAHAGNAFQDQ 205
DB 132 VVWVTHGGALVFGMASTLYDGSMTLALENVVVIYIYRLGVLDGFSTGDKHATGNMGYLDQ 191
QY 206 VVALSMVQKNIEFFGADPSVYTFGESAGATSVSSLISLSPMAKGLFHKAIMSGVAILIPY 265
DB 192 VVALRMVQONIAHFGNDRTYTFGESAGTSVSSLVSPISQGLFHGAIMSGVALLPBG 251
QY 266 LEADHYKSEDL--OVVAHFCGNNASDSEBALRCRTPKSKELLTLOKTKSFTRVVDGA 323
DB 252 LIA-----SSADVI STVVA NLSACDQVDSBALVGCIRGSKKEIILAINKPFKMI PCVVDGV 307
QY 324 FFPNEPDLISQAKAPKAPISITIGVNNHCGFLP-----MKEAPELISGSKSLAHLILQN 379
DB 308 FLPRHPOBILASADQPVPSIVGVNNNEFGWLI PKVMRIYDTQKEMDRASQALLOKMLT 367
QY 380 ILHIPPOYLILVANEYFHDKISLTERISDLIDLGDVFFVVPALITARYHRDAGAPVYFY 439
DB 368 LIMLPPTGDLREYIGDNGDPOTLOAQOFQEMMADSMFVLPALQVAHF-QCSRAPVYFY 426
QY 440 EFRHRPOCFEDTKPAFVADHADEVRFVFGAFLKGDIVMEGATEEBSKLSRKMMKTYA 499
DB 427 EFQHPSPWLKNI RPPHMKADHGDDEL PVPF-RSFCGNYIKF---TEEBEQLSRKMMKTYA 482
QY 500 TPARTEGNNGNDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDWTSTIP 550
DB 483 NFARNGNPNGBGLPMPPLFDQEBEYLOLNLQPAVGRALKAHRLQFWKXALP 533

```

Search completed: June 15, 2005, 09:46:09
Job time : 88 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 15, 2005, 09:39:56 ; Search time 866 Seconds

(without alignments)
3971.559 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 1 MPQGLTSSASQWCFLLIQP.....PLSLTFLSLIQPFFFCAP 581

Sequence:

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPFO.spool/6664091/runat_14062005.133640.10451/app_query.fasta.1.775
-DB=N Geneseq_16dec04 -QFWT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPRT=pco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=6664091.MCGN_1_1_470@runat_14062005.133640.10451 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=130 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16dec04:*

1: geneseqn1980b:*
2: geneseqn1990b:*
3: geneseqn2000b:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	2158	6	ABN84302
2	3079	100.0	2158	13	ADQ89093
3	2915	94.7	1728	6	ABQ86169
4	2900	94.2	2232	6	AAD40574
5	2878	93.5	2229	13	ACN42766

6	2877	93.4	2092	10	ADB22095
7	2607.5	84.7	1746	6	ABQ86170
8	2606	84.6	1962	10	ADAS3168
9	2597.5	84.4	1746	8	ABX72267
10	2588.5	84.1	1857	7	ADP19681
11	2015	65.4	2145	12	ADP50146
12	1985	64.5	1629	12	ADP50144
13	1780	57.8	1071	6	ABQ86171
14	1581	51.3	1244	10	ADCS5523
15	1230	39.9	1680	4	AAFP2558
16	1230	39.9	1680	8	AAFS2013
17	1230	39.9	1699	8	AB224052
18	1230	39.9	1748	8	AB224055
19	1230	39.9	2169	8	AB224051
20	1230	39.9	2191	6	ABL62491
21	1230	39.9	2191	6	ABL62492
22	1230	39.9	2191	6	ABL63768
23	1230	39.9	2191	6	ABN97359
24	1230	39.9	2484	3	AAE21884
25	1221	39.7	921	6	AAD33344
26	1215.5	39.5	1902	6	ABK62573
27	1215.5	39.5	1902	10	ADBS8928
28	1215.5	39.5	1902	10	ADBS3653
29	1215.5	39.5	1902	10	ABT40829
30	1215.5	39.5	1902	12	ADP71859
31	1191	38.7	2888	5	ABX71401
32	1181	38.4	2087	6	ABD24096
33	1181	38.4	2087	12	ADT79887
34	1179.5	38.3	2101	3	ADDO0680
35	1179.5	38.3	2275	13	ADSL0162
36	1179.5	38.3	3776	4	AAD08405
37	1179.5	38.3	3824	3	AAA96341
38	1179.5	38.3	3824	4	AAE21514
39	1179.5	38.3	3824	6	ABK33640
40	1179.5	38.3	3824	6	ABL88204
41	1179.5	38.3	3824	6	ABL95693
42	1179.5	38.3	3824	6	ABL95693
43	1179.5	38.3	3824	8	ACD28810
44	1179.5	38.3	3824	8	ACA03873
45	1179.5	38.3	3824	8	ABX89411

ALIGNMENTS

RESULT 1	ABN84302	standard; cDNA, 2158 BP.
ID	ABN84302	
XX	ABN84302;	
AC	23-SEP-2002	(first entry)
XX		
DE	Human carboxylesterase family member 53010 coding sequence.	
XX		
KW	Carboxylesterase; enzyme; human; analgesic; nocotropic; antiinflammatory; diagnosis; gene therapy; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	96..1841
FT		/*tag= a
FT		/product= "53010"
FT		/note= "the CDS is also claimed in Claim 1"
FT	sig_peptide	96..173
FT		/*tag= b
FT	mat_peptide	174..1838
FT		/*tag= c
XX		
XX	WO200250256-A2.	
XX		
XX	27-JUN-2002.	

PF 18-DEC-2001; 2001MO-US049075.
XX
PR 18-DEC-2000; 2000US-0256369P.
PR 28-MAR-2001; 2001US-0279508P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Siles-Santiago I;
XX
DR MPI; 2002-547936/58.
DR P-PSDB; ABB9537.

XX 53010 nucleic acids, useful for diagnosing and treating e.g. vascular
PT diseases, autoimmune diseases, or neurodegenerative diseases, as
PT surrogate markers, in tissue typing and chromosome mapping.
XX

PS Claim 1; Page 109-110; 122pp; English.

CC The present nucleotide sequence, the coding region of which is also
CC claimed, is that of cDNA encoding human 53010, a novel member of the
CC carboxylesterase family. The invention provides 53010 nucleic acids,
CC antisense molecules, expression vectors, host cells, transgenic animals,
CC 53010 proteins, fusion proteins, antigenic peptides, anti-53010
CC antibodies and methods for detecting the presence of 53010 polypeptides
CC or nucleic acids, of identifying a compound that binds to the 53010
CC polypeptide, and of modulating the activity of the polypeptide. The 53010
CC nucleic acids and polypeptides can act as novel diagnostic and
CC therapeutic agents for controlling disorders involving aberrant or
CC deficient hydrolysis of carboxylic esters. A 53010 mRNA is highly
CC expressed in the central and peripheral nervous system, and its
CC expression is regulated in some rodent pain models; 53010 molecules can
CC also act as novel diagnostic targets and therapeutic agents for
CC controlling neurological disorders, such as pain-related disorders. A
CC claimed method of treating or preventing a disorder (especially a pain-
CC related disorder) characterised by aberrant activity of a 53010-
CC expressing cell involves administering a compound that modulates 53010
CC activity or expression. 53010 nucleic acids are also useful in chromsome
CC mapping, tissue typing, in forensic biology, prognostic assays, in
CC arrays, for detection of variations or mutations, as surrogate markers
CC and in pharmacogenomics
XX

XX Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,71e-282 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-023-515-2 (1-581) x ABBN84302 (1-2158)

QY 1 MetProGlnGlyLeuThSerSerAlaSerGlnTrpCySPhePheLeuIleuGlnPro 20
Db 96 ATGGCACAGGAGACTTCACTTCATCTGCTTCACAAATGGTCTTTTCCCTCATCTCCAGCCC 155
QY 21 LeuLeuGlyHisArgGlnTrpGlyLeuThArgProSerAlaGlnGlyProGlnArgAsn 40
Db 156 CTGTTGGGACACACAGACTGGGGAACCTGGGCTTGTCTGAAGGGCCACAGAGAAC 215
QY 41 ThrArgLeuGlyTrpIleGlnGlyLeuGlnValThrValLeuGlySerProValProVal 60
Db 216 ACCAGGCTGGGATGATTCAGGGCAGACAGTCACTGTGGGAGAGCCCTGTCTGTG 275
QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
Db 276 AACGTGTTCTCTGGAGTCCCTTTGTGCTCCCGCTGGGATCCCTGGCAATTTACAAAC 335
QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnAlaThrSerTrpProAsnLeuCy 100
Db 336 CCGCAGCTGTGATGCCCTGTGGATTAATCTTGGAGAGAGCACTCTACCTTAATTGTGC 395

QY 101 LeuGlnAsnSerGlyTrpLeuLeuLeuAspGlnHisMetLeuLeuValHisTrpProLys 120
Db 396 CTCAGAACTCAGAGTGGCTGTCTTAGATCAACACATGCTCAAGTGCATTAACCCGAAA 455
QY 121 PheGlyValSerGlyAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
Db 456 TTGGAGTGTCAAGAGACTGCTCTTCACTTGAACATCTAGCCCTGCCAGCCGATAC 515
QY 141 GlySerTrpLeuProValLeuValTrpPheProGlyGlyAlaPheLysTrpGlySerAla 160
Db 516 GGCTCCAAAGTCCCGCTGTGTGTGTGTCCAGAGAGGTGCTTCAAGACTGGCTGACC 575
QY 161 SerLeuPheAspGlySerAlaLeuAlaAlaTrpGlyAspValLeuValValGln 180
Db 576 TCCATCTTGAATGGATCGCCCTGTGCTGCTATAGAGACGCTGTGTGTGTGTGTGTCAG 635
QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTrp 200
Db 636 TACCGGCTAGAGAAATTTTGGTTTCTTCAACCATGGGATTCAGCTCCGAGGAACTGG 695
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGluPheGly 220
Db 696 GCCTTCAAGGACACAGTGGCTGTCTGTGCTGGGTCCAGAAACATCGAGTCTTCGGT 755
QY 221 GlyAspProSerSerValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSer 240
Db 756 GGGAGCCCAAGCTTGTGACCATCTTGGGAGTCCGAGGAGCATTAAGTGTTCAGT 815
QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGlySerGlyVal 260
Db 816 CTATACGTGCTCCCATGCGCAAGGCTTATTCACAAACCATCATCAGAGAGTGGGTG 875
QY 261 AlaIleIleProTrpTrpLeuGlnAlaHisAspTrpGlyLysSerGlyAspLeuGlnVal 280
Db 876 GCCATCATCCCTTAACCTGAGGCCCATGATTATGAGAAAGTGGAGACTGTCAGAGTGGT 935
QY 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThr 300
Db 936 GCACATTTCTGTGTAAACATGCGTCAAGCTCTGAGGCCCTTGCTGAGTGTCTGAGGACA 995
QY 301 LysProSerLysGlyLeuLeuTrpLeuSerGlnLysThrLysSerPheThrArgValVal 320
Db 996 AAACCTCCCAAGGAGCTGTCAGCCCTCAGCCACAGAAACAAAGCTTTCACTCGAGTGGT 1055
QY 321 AspGlyValaPhePheProAsnGlnProLeuAspLeuLeuSerGlnLysAlaPheLysVal 340
Db 1056 GATGGTCTTCTTCTTCTTAATGAGCCCTAGATCTATGTCTCAGAAACATTTAAAGCA 1115
QY 341 IleProSerIleIleGlyValaAsnHisGlyCysGlyPheLeuLeuProMetLysGlu 360
Db 1116 ATTCTTCCATCATCGAGTCAATTAACACAGAGTGTGGCTTCTGCTGCTTATGAGAG 1175
QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIle 380
Db 1176 GCTCTGAGATCTCTCAAGTGGCTCCAAACAAAGTCCCTTGCCCTCATGATACAAACATC 1235
QY 381 LeuHisIleProProGlnTrpLeuHisLeuValAlaAsnGlnTrpPheHisAspLysHis 400
Db 1236 CTGCACATCCCGGCTCAGTATTTGCACTTGTGGCTTAAGTAATCTTCAATGCAAGCAC 1295
QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
Db 1296 TCCCTGACGTGAATTCGAGACAGTCTTCTGGAATCTTCTGGAATGTCTTTTGTGGTC 1355
QY 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGlu 440
Db 1356 CCTGACACTGATCAACAGCTGATTCACAGAGATGCTGGTGCACTGTCTAATCTTATGAG 1415
QY 441 PheArgHisArgProGlnCysPheGlyAspTrpLysProAlaPheValLysAlaAspHis 460
Db 1416 TTTCGGCACCAGGCTCAGTGTCTTTGAAAGACAGAACCCGAGCTTTGTCAAAGCCGACCAC 1475
QY 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480

QY	201	AlaPheIyIAepGIuValAlaIaleuSeRTpValGlnIySaenIIeGIuPhePheGIy	220
Db	696	GCCTTAAAGAACCAAGGTGGCTGCTGTCCTGGCTCCAGAAAGAACATGACTTCTTGCT	755
QY	221	GIyAspProSeSerSerValThrIIlePheGIyGInSerAlaGIyAlaIIeSerValSerSer	240
Db	756	GGGGAACCCAGCTCTGTGACCATCTTTGGCGAGTCCGGGAGGCCATAGGTTTCAGT	815
QY	241	IeuIIleIeuSeRProMeAlaIySGIyleuPheIIyIeValaIIeMeGIuSeRGIyVal	260
Db	816	CTTAAATCTCTCCATGGCCAAAGCTTATTCCCAAAAGCCATCATGGAGAGTGGGCTG	875
QY	261	AlaIIeIleProTYIeGIuIalAIIaSaPpTYGIyIySeSeRGIuAePleGIuVal	280
Db	876	GGCATCATCCCTTACTGAGGCCCATATTATAGAAAGATGAGACCTGCAGGTGGTT	935
QY	281	AlaHISpHeCYeGIyASnaSnaIaSeRAspSeRGIuAlaIeuIeAryCYeIaAqTTr	300
Db	936	GCACATTTCTGTGGTAAACATGGCGTCAACTGTAGGCCCTGCTGAGGTCTCCAGAGCA	995
QY	301	LySPProSeRTyBSGIuIeUleuThIeUSeRGIuIyThIySeSeRPhetHarGyVal	320
Db	996	AAACCTTCAGAGAGAGCTGACCTCAGCCAGCAAAACAAAGCTTTCATCTCAGAGGTT	1055
QY	321	AspGIyAlaPhePheProAsnGIuProIeUaSpIeUSeRGIuIySAlaPheIySAla	340
Db	1056	GATGGTCTTCTTCTTCCTAATGAGCCTCTGATCTATGTCTCAGAAAGCATTTAAAGCA	1115
QY	341	IIleProSeRIIeIeGIyValaIaSaPheHISGIyCyeGIyPheIeUeUProMeIySGIu	360
Db	1116	ATTCCTTCATCATTCGAGTCAATAAACAGAGATGGCTTCTGCTGCTCAATGAAGAG	1175
QY	361	AlaProGIuIleUSeRGIySeRAsnIySeSeIeUAlaIeUHIleUeIIeIaIaSnIe	380
Db	1176	GCTCTGAGATCCTCACTGAGCTCAACAAGTCCCTTGCCCTCATCTGATTAACAAATC	1235
QY	381	IeuHISIIeProProGIuTYIeUHIleUeValAlaSaGIuTYrPheHISaPlySHIS	400
Db	1236	CTGCACATCCCGGCTCAAGTATTTGACCTTGTGGCTAAAGTAATCTTCATGACAGAC	1295
QY	401	SeRleuThGIuIleAryAsPSeSeIeUaAsPleUeUeGIyAsPValPhePheVal	420
Db	1296	TCCCTGACTAAATCCGAGCAGCTCTTCTGACTTCTGGAGAGTGTGTTCTTGTGGTC	1355
QY	421	ProAlaIeUleThrIalAryTYrHISaTGAaPAlaGIyAlaProValTYrPheTYrGIu	440
Db	1356	CTCGACTGATCAACACTCGATATCAACAAGATGCTGGGACCTGTCTATCTTATAG	1415
QY	441	PheArHISaRProGIuCYSPheGIuAaPTrIySProAlaPheValIyAlaAsPHis	460
Db	1416	TTTCGGCACCGGCTCAGTGTCTTTGAAAGACAGAAACCGGCTTTGTCAAGCCGACAC	1475
QY	461	AlaAspGIuValAryPheValPheGIyGIyAlaPheIeUySGIyAsPAlaIeValMePhe	480
Db	1476	GCCTGAAGAGTCGCTTGTGTGTGGTGTGGCTTCTGGAAGGGGACATTGTTATGTC	1535
QY	481	GIuGIyAlaThGIuGIuIyIeUSeRGIyIySeMeTMeIyTYrTPAlaThr	500
Db	1536	GAAAGAGCCACGAGAGAGAGAGATTACTAGCCGGAAGAGAAATCTGGGGCTACC	1595
QY	501	PheAlaHqTTrGIyASnProAsnGIyASnaAsPleUSeRleUTrProAlaTYrASnIeU	520
Db	1596	TTTGTCTGAAACCGGGAATCTTAATGGAGCAACCTGTCTGTGGCCAGCTTAAATCTG	1655
QY	521	ThrGIuGIuTYIeUGIuIeUaAsPMeSeSeIeUeGIyGIuHqIeUySGIuPro	540
Db	1656	ACTGAGCATTACTTCACTGACTGTAACATGAGCCTCGAGACAGACCTCAAAAGAACCG	1715
QY	541	ArgValaAsPpHeTrPTrHSeRTrIIleProIeUleUSeRAlaSeRAsPMeTMeIyHIS	560
Db	1716	CGGGTGAATTTTGTGACAGACCATCCCTGTATCTGTCTCTCCATCATGCTTCAAC	1775
QY	561	SeRProIeUSeRleUThrPheUSeRleUeUInProPhePhePheCYaIa	580

Db 1776 AGCTCTCTTTCTTCTTAACCTTCCTCTCTCTCTCAGCCCTTCCTTTCTTTGTGCT 1835

QY 581 PRO 581
|||
Db 1836 CCT 1838

RESULT 3
AB086169
ID AB086169 standard; DNA, 1728 BP.
XX
XX AB086169;
AC
DT 10-SEP-2002 (first entry)
XX
XX Novel human gene. SEQ ID 40.
DE
XX
XX Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antinflammatory; cardiant; antilucer; vincide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo-transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200250105-A1.
PN
XX
XX 27-JUN-2002.
PD
XX
XX 17-DEC-2001; 2001WO-US049222.
PF
XX
XX 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
XX Agatwala P, Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX WPI. 2002-508784/54.
DR P-PsDB; ABP61004.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis;
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX Claim 2(a); Page 251, 335pp; English.
CC
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antinflammatory,
CC cardiant, antilucer, vincide, antithyroid, anti-infection;
CC

ID	AA040574	standard; cDNA; 2232 BP.
AC	AA040574;	
XX		
DT	30-OCT-2002	(first entry)
XX		
DE	Human drug metabolizing enzyme (DME-10) cDNA.	
XX		
KW	Human, drug metabolizing enzyme; autoimmune; inflammatory disorder;	
KW	acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;	
KW	proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;	
KW	asthma; neurological disorder; Alzheimer's disease; Huntington's disease;	
KW	dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;	
KW	drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;	
KW	renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;	
KW	anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;	
KW	goltze; gastrointestinal disorder; gene therapy; virucide; anticonagulant;	
KW	anticonvulsant; nootropic; enzyme; DME-10; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1929
FT		/*tag= a
FT		/product= "Human DME-10"
FT	sig_peptide	1..150
FT		/*tag= b
FT	mat_peptide	151..1926
FT		/*tag= c
FT		/product= "Mature human DME-10"
PN	MO200246426-A2.	
PD		
XX	13-JUN-2002.	
PF		
XX	04-DEC-2001; 2001MO-US047429.	
PR	08-DEC-2000; 2000US-0254308P.	
PR	15-DEC-2000; 2000US-0256189P.	
PR	21-DEC-2000; 2000US-0257713P.	
PR	19-JAN-2001; 2001US-0262706P.	
XX	02-FEB-2001; 2001US-0266020P.	
XX		
XX	(INCY-) INCYTE GENOMICS INC.	
PI	Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ;	
PI	Lee EA, Ding L, Hafalia AJA, Tang YT, Yue H, Tribouley CM, Lu DM;	
PI	Lai PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y,	
PI	Ison CH;	
XX		
DR	WPI; 2002-519668/55.	
XX		
XX	P-PSDB; AAE25025.	
PT	Novel human drug metabolizing polypeptide, useful in diagnosis,	
PT	prevention or treatment of autoimmune/inflammatory, cell proliferative,	
PT	neurological, developmental, endocrine, metabolic and gastrointestinal	
PT	disorders.	
XX		
XX	Claim 78; Page 167-168; 169pp; English.	
XX		
XX	The invention relates to an isolated human drug metabolizing enzyme (DME	
CC	and its nucleotide. DME is useful for diagnosing, treating or preventing	
CC	disorders associated with aberrant expression of DME, where the disorders	
CC	are selected from autoimmune/inflammatory disorder such as acquired	
CC	immunodeficiency syndrome (AIDS), asthma, atherosclerosis, psoriasis,	
CC	infects; a cell proliferative disorder such as arteriosclerosis,	
CC	cirrhosis, hepatitis, and cancer; a neurological disorder such as	
CC	Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;	
CC	a developmental disorder such as renal tubular acidosis, epilepsy,	
CC	anaemia; an endocrine disorder such as adenoma, thrombosis and infections	
CC	; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic	
CC	disorder such as cystic fibrosis, diabetes and goltze; a gastrointestinal	

```

QY 291 SerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer 310
DB 1054 TCTGAGGCGCTCTGAGGTCCTGAGAGCAAAACCTCCAGAGAGCTGTCAGCCCTCAGC 1113
QY 311 GlnLysThrLysSerPheThrArgValValAspGlyValaPhePheProAsnGluProLeu 330
DB 1114 CAGAAACCAAAAGCTTTCACCTCGAGTGGTGAATGGCTTTCTTCTTCTTAATAGGCTCTTA 1173
QY 331 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValaAsnAsnHis 350
DB 1174 GATCTATTGCTCAGAAACCAATTAAGCAATTCCTTCATCATCGAGTCAATAAACCAAC 1233
QY 351 GluGlyGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerLysSerLys 370
DB 1234 GAGTGTGGCTTCTGCTGCTTATGAAAGAGGCTCTGAGGGTCTCTAGTGGCTCCACAAAG 1293
QY 371 SerLeuAlaLeuHisAlaLeuIleGlnAsnIleLeuHisIleProProGlnIleLysHisLeu 390
DB 1294 TCCCTTGGCCCTCATCTGATACAAACATCCCGACATCCCGCTCAGTATTTGGACCTT 1353
QY 391 ValAlaAsnGluLysPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeu 410
DB 1354 GTGGCTAAATGAATACTTCCATGACAAAGCACTCCCTGACCTGAATCCGAGACAGTCTTCTG 1413
QY 411 AspLeuLeuGlyAspValaPhePheValValProAlaLeuIleThrAlaArgThrHisArg 430
DB 1414 GACTTGCTTGGAGATGATGTTCTTTGTGCTCTGACCTATCATCAGCTGATATACAA 1473
QY 431 AspAlaGlyAlaProValLysPheLysArgLysPheArgHisArgProGlnCysPheGluAsp 450
DB 1474 GATGCTGGGACACCTGTACTTCTATGATTTGGGACCGGCTCAGTGGCTTTGAAAGAC 1533
QY 451 ThrLysProAlaPheValLysAlaAspHisAlaAspGluValaArgPheValaPheGlyGly 470
DB 1534 AGGAAAGCGGCTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGTGCTGGTGGT 1593
QY 471 AlaPheLeuLysGlyAspIleValaMetPheGluGlyAlaThrGluGluGluLysLeuLeu 490
DB 1594 GCTTCTCGAAGGGGACATGTATGATGTCAGAGGACCAAGAGAGGAGAAAGTTACTG 1653
QY 491 SerArgLysMetLeuLysLysTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
DB 1654 AGCGGAGATGATGAATATCTGGGTCCTTGTGCTCGAAGCGGGAATCTTAATGGAAAC 1713
QY 511 AspLeuSerLeuTrpProAlaThrAsnLeuThrGluGlnIleLysLeuAsnLeuAsn 530
DB 1714 GACCTGTTTCTGTGGCCAGCTTATATCTGACTGAGACGATACCTCCAGCTGAGCTTGAAC 1773
QY 531 MetSerLeuGlyGlnArgLeuLysGluProArgValaAspPheTrpThrSerThrIlePro 550
DB 1774 ATGAGCCCTCGACAGACTCAAGAAACCGCGGAGGATTTTGGACCGACCAACATCCCC 1833
QY 551 LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 570
DB 1834 CTGATCTCTGTCTGCTCCGACATGCTCCACATCTCTTCTTCTTAATCTTCTCTCT 1893
QY 571 LeuLeuGlnProPhePhePhePheCysAlaPro 581
DB 1894 CTCTCCAGCCTTCTTTTCTTTGTGCTCT 1926

```

RESULT 5

ACN42766
ID ACN42766 standard; cDNA; 2229 BP.

ACN42766;
AC ACN42766;
XX 18-NOV-2004 (first entry)

XX Human diagnostic and therapeutic polynucleotide SEQ ID NO:1641.
DB Human diagnostic and therapeutic polynucleotide; 89; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX dithp.

XX 89; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
XX dithp.

```

XX OS Homo sapiens.
XX PN MO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCYTE) INCYTE CORP.
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
PI Hartschornke TM, Suchorolski MT, Altus CM, Plets SJ, Elder LV,
PI Mooney EM, Delegeane AM, Panesar IS, Banyille SC, Reddy TP,
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Geregin EH,
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL,
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vilt UA, Kilton BS,
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D,
PI Patury S, Shi X, Suarez CO;
XX WPI; 2004-329368/30.
DR P-PSDB; ABW84114.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PS in gene mapping.
XX Claim 1, Page: 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorder, developmental disorder, endocrine
XX disorder, neurological disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germ-line
XX gene therapy. The present sequence represents a dithp polynucleotide of
XX the invention. Note: The sequence data for this patent is not represented
XX in the printed specification, but was obtained in electronic format
XX directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
SQ Sequence 2229 BP; 503 A; 585 C; 547 G; 594 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6,17e-263 Length: 2229
Score: 2878.00 Matches: 549
Percent Similarity: 99.82% Conservative: 1
Beet Local Similarity: 99.64% Mismatches: 1
Query Match: 93.47% Indels: 1
DB: Query Match: 13 Gaps: 0
US-10-023-515-2 (1-581) x ACN42766 (1-2229)
QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyGln 50
DB 274 GGGCCCTTCTGCTGAAAGGCGACAGAGAAACACAGGCTGGATGATTCAGAGGCAAGCAA 333
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValaProPheAlaIle 70
DB 334 GTCACTGTGCTGGAGAGCCCTGTGCTGAAACGTTTCTCTGAGTCCCTTCTTCTGCT 393
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 90
DB 394 CCCCCTGGGATCCCTCGATTTACGAACCGGACGCTGATCGCCCTGGGATTAACCTTG 453

```


QY	91	ArgGluAlaIeThrSerTyrProAsnLeuCySeuGlnAsnSerGluTrpLeuLeuAsp	110
Db	454	CGAAGAACCACTCTCAACCTTAATTGGCCCTCGAAGACTCAGAGTGGCTGCTTAGAT	513
QY	111	GlnHsMeLeuLeuValHisTyrProIysPheGlyValSerGluAspCySeuTyrLeu	130
Db	514	CAACATATGCTCAAGAGTGCATTACCGAAATTCGAGGTGTAAGAAGACTGCTCTACCTG	573
QY	131	AsnIleTyrAlaProAlaHisIleAsnThrGlySerIysLeuProValLeuValTrpPhe	150
Db	574	AACATCTATGGCGCCGCCCAAGCATACAGAGCTCAAGCTCCCGCTTGGTGGTTC	633
QY	151	ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAla	170
Db	634	CCAGAGAGTGGCTTCMAAGCTGGCTCAAGCTCCACTTATATGGGTCCGGCCGGCTGGCC	693
QY	171	TyrGluAspValLeuValValGlnTyrArgLeuGlyIlePheGlyPhePheThr	190
Db	694	TATAGAGACGTGGCTGGTGTGGTGTGCTCAGTACCGGCTAGAGATATTGGTTCTC-ACG	752
QY	191	ThrTrpAspGlnHisAlaProGlyAsnThrAlaPheLeuAspGlnValAlaLeuSer	210
Db	753	ACATGGATTCAGCATGCTCCGGGAACTGGGCTTCAAGAACCAAGGTGGTGTCTGTCC	812
QY	211	TrpValGlnIysAsnIleGluPheGlyGlyAspProSerSerValThrIlePheGly	230
Db	813	TGGGTCCAGAAACATCGAGTTCCTGGTGGGACCCAGCTGTGACATCTTTGGC	872
QY	231	GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu	250
Db	873	GAGTCCCGGAGGCCATTAAGTGTCTTAGTCTTATCTGTCTCCATGGCCAAAGCTTA	932
QY	251	PheHisValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAsp	270
Db	933	TTCCACAAACCATCATCGAGAGTGGGGTGGCCATCATCTCTTACCTGGAGCCCATAT	992
QY	271	TyrGluIysSerGluAspLeuGlnValValAlaHisPheCyGlyAsnAsnAlaSerAsp	290
Db	993	TATGAGAAAGTGAAGACCTTGCAGGGGTGTGACATTTGTGGTAAACATCCGCAAC	1052
QY	291	SerGluAlaLeuLeuArgCySeuAsnArgThrIysProSerIysGluIleLeuThrLeuSer	310
Db	1053	TCGAGAGCCCTGTGAGTGGCTGAGAGCAAAACCTCCAAAGAGCTGTACCTCAGC	1112
QY	311	GlnIysThrIysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeu	330
Db	1113	CGAAGAACAAAGCTTTCACCTGAGAGGTGTGAATGGCTTCTTCTTAAGAAGCTCTA	1172
QY	331	AspLeuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAsnAsnHis	350
Db	1173	GATCTATTGTCTCAGAAACATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC	1232
QY	351	GluCyGlyPheLeuLeuProMetIysGlyAlaProGluIleLeuSerGlySerAsnIys	370
Db	1233	GAGTGTGGCTTCTGTGCTTGAAGAGGCTCTTGAGATCTCAAGTGGCTCCAAAG	1292
QY	371	SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisLeu	390
Db	1293	TCCCTTGGCCCTCATCTGATACAAACATCTCGACATCCGGCTCAGTATTGGACCTT	1352
QY	391	ValAlaAsnGluTyrPheHisAspIysHisSerLeuThrGluIleIleAspSerLeuLeu	410
Db	1353	GTGGCTAATGAATCTTCAATGACAAAGACTCCCTGATGTAAATCCGACAGCTTCTCG	1412
QY	411	AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArg	430
Db	1413	GACTTGCTGAGATGTGTCTTGTGGTCCCTGACATGATCAAGTCCATATCACAA	1472
QY	431	AspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnIysPheGluAsp	450
Db	1473	GATGCTGTGACCTGTCTACTCTTAAGATTTTGCGACCGGCTCAGTCTTTGAAGAC	1532

Oy		451	ThrlYsrProAlaPheValIylValaaSPHsaIAaSPGIuValArgPheValPheGlygly	470
Db		1533	ACGAAGCCACTTTTGTCAAAAGCCGACAACGGCTGATGAAGTCGCCCTTGTTGTCGTGCT	1592
Oy		471	AlaPheLeuLysGlyAspIleValMetPheGluGlyValAthrGluGluGluLeuLeu	490
Db		1593	GCCCTCCTCGAAGGGGACATTGTTANTGTTGAAAGACCACGACGAGAGAGAACTTACTG	1652
Oy		491	SerArgLysMetMetLysTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn	510
Db		1653	AGCCGAGAAGTATGAATAATCTGGGCTACTTGTGCTCGAACCCGGAATCCTTAATGGGAA	1712
Oy		511	AspLeuSerLeuTrpProAlaTyrrAmLeuThrGluGlnTyrrLeuGlnLeuAspLeuAsn	530
Db		1713	GACTGCTCTCTGTGGCACCTTAATATCGACTGAGAGTACCTCCAGCTGACTTGAAAC	1772
Oy		531	MetSerLeuGlyValArgLeuLysGluProArgValaAspPheTrpHisSerThrIlePro	550
Db		1773	ATGAGCTCGGACAGAGACTCAAAGAACCCGGGGTGATTTTTGGACCAAGCATCCCC	1832
Oy		551	LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer	570
Db		1833	CTGATCTGTCTGCCTCCGACATGCTCCACAGTCCCTCTTCTTCTTAACCTTCTCT	1892
Oy		571	LeuLeuGlnProPhePhePhePheCysAlaPro	581
Db		1893	CTCCTCCAGCCTTCTTCTTCTTCTTGTGCTCTCT	1925
RESULT 6				
ADBB62095				
ID	ADB62095	standard; cDNA; 2092 BP.		
XX				
AC	ADB62095;			
XX				
DT				
DE	04-DEC-2003	(first entry)		
XX				
Human	cDNA encoding clone BRAWH20021910.			
XX				
KW	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;			
KM	tissue regeneration; cell regeneration; membrane protein;			
KW	signal transduction-related protein; transcription-related protein;			
KM	osteoporosis; neurological disease; cancer; tumour.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	394..1803		
FT		/*tag= a		
FT		/product= "Clone BRAWH20021910 protein"		
FN	EPI308459-A2.			
PD				
PD	07-MAY-2003.			
PF				
PF	28-MAR-2002; 2002EP-00007401.			
PR				
PR	05-NOV-2001; 2001JP-00379298.			
FR	25-JAN-2002; 2002US-00350978.			
PA				
PA	(HELI-) HELIX RES INST.			
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.			
Pt	Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehii S;			
Pt	Yanamoto J, Isono Y, Hiro Y, Otsuka K, Nagai K, Irie R, Tamechika I;			
Pt	Seki N, Yoshikawa T, Otsuka M, Nagahashi K, Maezono Y;			
DR	WPI: 2003-450961/43.			
DR	P-PsDB: ADB64065.			
XX				
XX	New polynucleotides and polypeptides, useful for developing a diagnostic			
XX	marker or medicines for regulation of their expression and activity, or			
XX	as targets of gene therapy.			

PS Claim 1; Page: 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.

XX Sequence 2092 BP; 464 A; 552 C; 497 G; 579 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7e-263	Length:	2092
Score:	2877.00	Matches:	551
Percent Similarity:	98.40%	Conservative:	2
Best Local Similarity:	98.04%	Mismatches:	2
Query Match:	93.44%	Indels:	7
DB:	10	Gaps:	2

US-10-023-515-2 (1-581) × ADB62095 (1-2092)

QY	24	HISAAGGINTTPRGlylyThrGlyProSerAlaGluGlyProGlnArgAsnThrArgLeu	43
DB	126	CACCTTAACTGG-----GGGCTTCTGCTGAAGGGCCACAGAGAACCGAGGCTG	176
QY	44	GLYTPRIeGlnGlyLeuGlnValThrValLeuGlySerProValProValAsnValPhe	63
DB	177	GGATGGATTCCAGGAGCAAGCACTGCTGCTGGAGAGCCCTGCTGTGAACGTGTTTC	236
QY	64	LeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsnProGlnPro	83
DB	237	CTCGGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGCGATTACGAAACCGGAGCTT	296
QY	84	AlaSerProTTPAspAsnLeuArgGluAlaThrSerTyrProAsnLeu-----Cyl	100
DB	297	GCATGGCCCTGGGATTAACCTTGCGAAGAGCCACTCTTACCTTAATTGTAAAGACAGGTG	356
QY	100	AlaLeuGlnAsnSerGlyUTTPLeuLeuLeuAspGlnHisMetLeuLysValHisTyrProly	120
DB	357	CCTCAGAACTCAGAGTGAGCTGCTCTTATGATCAACAATGCTCAAGGTGATTAACCGAA	416
QY	120	spheGlyValSerGluAspCylLeuTyrLeuAsnIleTyrAlaProAlaHisAlaAspThr	140
DB	417	ATTGGAGAGTGCAGAAAGACTGCTTCACTGAACATCATATGCGCTGCGCCAGCGGATAC	476
QY	140	rglySerlyLeuProValLeuValITTPheProGlyGlyAlaPheLysGlySerAla	160
DB	477	AGGCTCCAGAGCTCCCGTCTGTGTGTGTCCAGAGAGTGCTTCAAGACTGCGTCAAGC	536
QY	160	AserIlePheAspGlySerAlaLeuAlaAlaTyrGlnAspValLeuAlaValAlaGln	180
DB	537	CTCCATCTTTATAGGTGCGCCCTGGCTGCTTAAGAGACGTGCTGTGTGTGTGTCTCA	596
QY	180	nTyrArgLeuGlyIlePheGlyPhePheThrThrTPAspGlnHisAlaProGlyAsnTr	200
DB	597	GTACCGGCTAGGAATATTGTTTCTTCAACAATGGATGATGATGCTCGGGGAACTG	656

QY	200	PalAspPheLysAspGlnValAlaAlaLeuSerTyrValGlnLysAsnIleGluPhePheGln	220
DB	657	GGCTTCAAGAGCAAGAGTGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	716
QY	220	rglyAspProSerSerValITTPheGlyGluSerAlaGlyAlaIleSerValSerSe	240
DB	717	TGGGAGCCCAAGCTGTGAGCAATCTTTGGCGAGTCCCGGGAGCCATAAGTGTCTTAG	776
QY	240	rLeuIleLeuSerProMetAlaLysGlyLeuPheHisIleValIleMetGluSerGlyVa	260
DB	777	TCTTATAGTCTTCCCAAGGCGMAAGGCTTATTCACAAAGCATCATGAGAGTGGGT	836
QY	260	AlaIleIleProTyrLeuGlnValHisAspTyrGlnLysSerGluAspLeuGlnValVa	280
DB	837	GGCCATATCCCTTACCTGAGAGCCCAATATTAAGAGTAGAGAGCTGCAAGTGTGT	896
QY	280	AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCylLeuArgTh	300
DB	897	TGCACATTTCTGTGTAAACAATGCTGACACTGAGGCCCTGCTGAGGTGCTGAGGAC	956
QY	300	rlyAspProSerlyGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVa	320
DB	957	AAACCCCTCCAGAGAGCTGAGCCCTCAGCCAGAAACAAAGTCTTCACTGAGTGT	1016
QY	320	lAspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAl	340
DB	1017	TGATGAGCTTCTTCTTCTTAATAGCTCTTATGATCTTATGCTTCAAGAAAGCATTTAAAC	1076
QY	340	AlleProSerIleIleGlyValAsnAsnHisGluCylGlyPheLeuLeuProMetLysGln	360
DB	1077	AATTCCTTCCATCANTCGGAGTCAATATGACAGAGTGGCTTCCGTGCTGCTATAGGA	1136
QY	360	uAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnI	380
DB	1137	GGCTCTGAGATCTCCAGTGTGCTCCAAAGTCCCTTCCCTCCATCGATACAAACAT	1196
QY	380	eLeuHisIleProProGlnITyrLeuHisLeuValAlaAsnGluTyrPheHisAspLysHi	400
DB	1197	CTTGCAATCCCGCTCAGATTTTGACATTTGTGCTGCTTAATGATCTTCCATGACAGCA	1256
QY	400	sSerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlnIleAspValIlePheValVa	420
DB	1257	CTCCCTGACTGAATCCGAGACAGTCTTCTGGAAGTGTGCTTGTGTGTGT	1316
QY	420	lProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGln	440
DB	1317	CCCTGCACTGATCAGACTCATATACAGAGATGCTGTGTCACCTGTCTACTTATAGA	1376
QY	440	uPheArgHisArgProGlnCylspheGluAspThrLysAspProAlaPheValLysAlaAspHi	460
DB	1377	GTTCGGCAACCGGCTCAGTGTGTTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACCA	1436
QY	460	AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValIleMetPh	480
DB	1437	CGCTATATAGTCCGCTTGT	1496
QY	480	eGluGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTyrTTPAlaTh	500
DB	1497	CGAAGAGCCACGAGAGAGAGAGTACTGAGCCGGAAGATGATGAAGAAATCTGGGCTAC	1556
QY	500	rPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTTPProAlaTyrAsnLe	520
DB	1557	CTTGTGCTGAACCGGGAATCTTAATGGGAGAGCACTGTCTCTGTGGCAGCTTATTAATCT	1616
QY	520	uThrGlnGlnITyrLeuGlnLeuAspLeuAsnMetSerLeuGlnIleArgLeuLysGlnPr	540
DB	1617	GACTAGAGATACCTGAGCTGAGCTTGAACATAGAGCTCCGAGACAGACTCAAGAAACC	1676
QY	540	oArgValAspPheTTPThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHi	560
DB	1677	GCGGGTGAAGTGTGGACACACACATCCCTGATCTGTCTGCTGCTCGACATGCTCCA	1736

QY 560 sSerProLeuSerLeuThrPheLeuSerLeuGlnProPhePhePheGSA1 580
 DB 1737 CAGTCCCTCTTCTCTTACCTTCTCTCTCTCCAGCTTCTTCTTCTTGTGC 1796
 QY 580 APRO 581
 DB 1797 TCCCT 1800
 RESULT 7
 ABQ86170
 ID ABQ86170 standard; DNA, 1746 BP.
 XX ABQ86170;
 AC
 XX 10-SEP-2002 (first entry)
 DT
 XX
 DE Novel human gene. SEQ ID 41.
 KW Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
 KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
 KW antiinflammatory; cardiac; antitumor; virucide; antithyroid;
 KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
 KW wound healing disorders; atherosclerosis; Parkinson's disease;
 KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 KW inflammation; neoplastic disease; nervous system disorder;
 KW cardiovascular disorders; pancreatitis; respiratory disorder;
 KW hyperproliferation; systemic autoimmunity; hyper-immunity;
 KW developmental abnormality; gastrointestinal ulceration; neuropathy;
 KW haematological disease; metabolic disease; sperm dysfunction;
 KW thyroid disorder; hypothyroidism; brain damage; colitis;
 KW cone photo- transduction deficiency; neurological disease; stroke;
 KW anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
 KW growth abnormality; precocious puberty; gene; ss.
 XX Homo sapiens.
 OS
 XX
 PN WO200250105-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 17-DEC-2001; 2001WO-US049232.
 XX
 PR 19-DEC-2000; 2000US-0256710P.
 PR 20-DEC-2000; 2000US-0257048P.
 PR 09-JAN-2001; 2001US-0260482P.
 PR 30-JAN-2001; 2001US-0264922P.
 PR 06-FEB-2001; 2001US-0266797P.
 PR 19-MAR-2001; 2001US-0276988P.
 PR 04-APR-2001; 2001US-0281535P.
 PR 08-MAY-2001; 2001US-0289622P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Agarwal P, Birkeland M, Cogswell JP, Kahnack KF, Lai Y,
 PI Mathiesen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
 XX
 DR WPI; 2002-508784/54.
 DR P-PSDB; ABP61005.
 XX
 PT Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX
 PS Claim 2(a); Page 251-252; 335pp; English.
 CC The invention relates to an isolated polypeptide with signal sequences
 CC which allow it to be secreted extracellularly or membrane associated. The
 CC activity of polypeptides of the invention may be described as,
 CC cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,

CC cardiac, antitumor, virucide, antithyroid, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder,
 CC hyperproliferation, systemic autoimmunity, hyper-immunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic disease, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, anglogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
 CC of the invention
 XX

SO Sequence 1746 BP, 401 A; 464 C; 459 G; 422 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,17e-237 Length: 1746
 Score: 2607.50 Matches: 499
 Percent Similarity: 92.67% Conservative: 7
 Best Local Similarity: 91.39% Mismatches: 25
 Query Match: 84.69% Indels: 15
 DB: Gaps: 3

US-10-023-515-2 (1-581) x ABQ86170 (1-1746)

QY 12 TrpCysPhePheLeuLeuGlnProLeuLeuGlnYHisArgGlnTrpGlyThrGly 31
 DB 40 TGGGCTATCTGGGGTCTCTGACGCCCC-----ACCAAGGCG 75
 QY 32 ProSerAlaGlnGlyProGlnArgAenThrArgLeuGlyTrpIleGlnGlySerGlnVal 51
 DB 76 CCTTCTGTGAAGGCGCACAGAGAACCCAGGCTGGATTCGATTCAGGCGCAAGATTC 135
 QY 52 ThrValLeuGlySerProValProValAanValPheLeuGlyValProPheA1a1aPro 71
 DB 136 ACTGTGCTGGAGGCCCTGTGCTGGAACGTGTTCTCGAGATCCCTTCTGCTGCC 195
 QY 72 ProLeuGlySerLeuArgPheThrAenProGlnProAlaSerProThrAenLeuArg 91
 DB 196 CCGCTGGAGTCCCTGCGATTACGAACCCGACGCTGCATCCCGCTGGATTAATTCGCA 255
 QY 92 GluAlaThrSerTrpProAenLeuGlnAenSerGluTrpLeuLeuAanPgin 111
 DB 256 GAAGCCACCTCCCTAATTTGTGCTCCAGAACTCAGATGGCTGCTTAATCA 315
 QY 112 HisMetLeuValHisTrpProLysPheGlyValSerGluAanPgyLeuTrpLeuAan 131
 DB 316 CATATCTCAAGGTGATTAACCGAAATTCGAGGTGCAAGACTGCTCACTGAAC 375
 QY 132 IleTrpAlaProAlaHisAlaAenThrGlySerTrpLeuProValLeuValTrpPhePro 151
 DB 376 ATCTATGCGCTGCCAGCGCATACAGGCTCAAGCTCCCGCTTGTGTGTGTTCCCA 435
 QY 152 GlyValAlaPheLysThrGlySerAlaSerIlePheAanPgySerAlaLeuAlaTrp 171
 DB 436 GAGAGGCTTCAACACGTGCTCAGCTTCATCTTGTATGGGTCCGCTGCTGCTCAT 495
 QY 172 GluAanPheValValValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrThr 191
 DB 496 GAGGAGCGT 555
 QY 192 TrpAanPheHisAlaProGlyAanTrpAlaPheLysAanPginAlaAlaLeuSerTrp 211
 DB 556 TGGGATTCAGCATGTCTCCGGGAACTGGGCTTCAAGGACGAGTGGCTGTCTCTCGG 615
 QY 212 ValGlnLysAanIleGluPhePheGlyGlyAanProSerSerValThrIlePheGlyGlu 231
 DB 616 GTCCAGAAAGAACATCGAGTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 675

Db	405	GAAGCCACCTCTTACCCTTAATTGTGCTCCAGAACTTGAAGTGGCTGCTTATGATCAA	464
QY	112	HisMetLeuIysValHisIleTyrProIlysbPheGlyValSerGluAapCysLeuTyrLeuAsn	131
Db	465	CACATGCTCAAGAGTGCATTTACCCGAAATTCCGAGGTGTCAAGAAAGCTGCTTACCTTAAC	524
QY	132	IleTyrAlaProAlaHisAlaAspThrGlySerTyrLeuProValLeuValTrpPhePro	151
Db	555	ATCTATGGCCCTGCCCAAGCCGAAACAGAGCTCCAAAGCTCCCGCTTGGTGGTGTCCCA	584
QY	152	GlyGlyAlaPheIysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleTyr	171
Db	585	GGAGGTGGCTTCAGAGACTGGCTCAGGCTCCATCTTTATGGGTCCGCGCTGCTCTAT	644
QY	172	GluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr	191
Db	645	GAGAGCCGTGCTGGTGTGGTGTGTGTGCAGTACCGGCTAGAAATTTGGTGTCTTCAACA	704
QY	192	TrpAspGlnHisAlaProGlyValAsnThrAlaPheIysAspGlnValAlaAlaLeuSerTrp	211
Db	705	TGGGATCAGCATGCTGTCCGGGAACTGGAGCTTCAGAGACAGAGGTGGTGTGCTGTGG	764
QY	212	ValGlnIysAsnIleGluPhePheGlyGlyIysAspProSerSerValThrIlePheGlyGlu	231
Db	765	GTCAGAGAAACATCGAGTCTTGCGTGGGGACCCAGCTGTGTGACATCTTTGGCGAG	824
QY	232	SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeuPhe	251
Db	825	TCCGCGGAGACCATATGATGTTTCTTAGTCTTAACTGTCTCCCAATGGCCAAAGGCTTAATTC	884
QY	252	HisValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyr	271
Db	885	CACAAAGCCATCATGGAGAGTGGGGTGGCCATCATCTCCATTACCTGGAGGCCATATTTT	944
QY	272	GluIysSerGluAspLeuGlnValValAlaHisPheCysGlyIysAsnAlaSerAspSer	291
Db	945	GAGAAAGTGAAGACCTCGCAGGTGTGTGACATTTGTGGTAAACATCGTCAAGCTCT	1004
QY	292	GluAlaLeuLeuIysArgCysLeuAlaGlnTyrIlyProSerTyrGluIleLeuThrLeuSerGln	311
Db	1005	GAGGCCCTGTGAGTGGTGCAGAGCAAAACCTCCAAAGAGCTGTGACCTCCAGCCAG	1064
QY	312	IysThrIlySerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp	331
Db	1065	AAAAACAAGCTTTTCACTGAGAGGTGTGATGGCTTTCTTCTTAATGAGCCTTAAGT	1124
QY	332	LeuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAsnAsnHisGln	351
Db	1125	CTATTGTCTCAGAAAGCATTTAAAGCAATCTTCTCATTTATGGAGTCAATACCAAGAG	1184
QY	352	CysGlyPheLeuLeuProMetIysGluAlaProGluIleLeuSerGlySerAsnIysSer	371
Db	1185	TGTGGCTTCTGTGCTGATGAAAGAGGCTCTGTGAGATCTCAGTGGCTCCAAAGTCC	1244
QY	372	LeuAlaLeuHisIleuIleGlnAsnIleLeuHisGluIleProGluIlyLeuHisIleuVal	391
Db	1245	CTTGCCCTCACTGATACAAACATCTCTGACATCCGAGCTCCAAATATTTGACCTTGTG	1304
QY	392	AlaAsnGluTyrPheHisAspIlyHisSerIleuThrGluIleArgAspSerLeuLeuAsp	411
Db	1305	GCTAATGATATCTTCCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTTCTGAC	1364
QY	412	LeuLeuGluIysAspValPhePheValValAlaProAlaLeuIleThrAlaArgTyrHisAspAsp	431
Db	1365	TTGCTTGGAGATGATGTCTTGTGGTCTCCTGACATGACAGAGCTCGATATCAACAGA---	1421
QY	432	AlaGlyAlaProValTyrPheTyrGluPheAsnHisArgProGlnCysPheGluAspThr	451
Db	1421	-----	1421
QY	452	LysProAlaPheValIysAlaAspHisValaAspGluValArgPheValPheGlyGlyAla	471
Db	1421	-----	1421

Qy	472	PheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuSer	491
Db	1422	-----GAGAGACGACGAGGAGAGAGAGTACTGAGC	1454
Qy	492	ArgLysMetMetLysTyrTrpAlaThrPheAlaArgThnGlyAaspProAngLYAasnAsp	511
Db	1455	CGGAAGATCATGAAATACTGGGCTACCTTTGCTGCACC GGGAATCTTAATGGAAACGAC	1514
Qy	512	IeuSerLeuTrpProAlaTyrAsnLeuThnGluGlnTyrLeuGlnLeuAspLeuAsnMet	531
Db	1515	CTGTCTCTGTGGCCGACTTAATTCGTACAGCAAGTACCTCCGACCTGACTTGAAATCATG	1574
Qy	532	SerLeuGLYglnArgLeuLysGLUProArgValAspPheTrpTrpSerThrIleProLeu	551
Db	1575	AACCTCGACAGACGACTAAAGAACCGGGGTGATTTTTTGACACACACATCCCCCTG	1634
Qy	552	IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThnRheLeuSerLeu	571
Db	1635	ATCCCTGTGTGCTCGACATGCTCCACAGTCTCTTTCTTCTTAATTCTCTCTCTC	1694
Qy	572	LeuGlnProPhePhePhePheCysAlaPro	581
Db	1695	CTCCAGCCTTCTTTTCTTTTGTGCTCTCT	1724
RESULT 9			
ABK72267			
ID	ABX72267	standard; cDNA; 1746 BP.	
AC	ABX72267;		
XX			
XX			
DT	03-JUN-2003	(first entry)	
DE			
Xx	Human NOVX polynucleotide #98.		
Kw	Human; NOvX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;		
Kw	hypertension; congenital heart defect; aortic stenosis; valve disease;		
Kw	atrial septal defect; atrioventricular canal defect; ductus arteriosus;		
Kw	pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;		
Kw	tuberos sclerosiis; acroderma; attherosclerosis; infectious disease;		
Kw	obesity; anoexia; neurodegenerative disorder; Alzheimer's disease;		
Kw	Parkinson's disease; immune disorder; haematopoietic disorder;		
Xx	haemophilia; hypercoagulation; Crohn's disease; cancer.		
OS	Homo sapiens.		
XX			
XX			
PN	MO200281498-A2.		
PD			
XX	17-OCT-2002.		
XX			
PF	03-APR-2002; 2002MO-USO10780.		
XX			
PR	03-APR-2001; 2001US-0281086P.		
PR	03-APR-2001; 2001US-0281136P.		
PR	05-APR-2001; 2001US-0281863P.		
PR	06-APR-2001; 2001US-0281906P.		
PR	10-APR-2001; 2001US-0282020P.		
PR	10-APR-2001; 2001US-0282930P.		
PR	10-APR-2001; 2001US-0282934P.		
PR	12-APR-2001; 2001US-0283512P.		
PR	13-APR-2001; 2001US-0283710P.		
PR	17-APR-2001; 2001US-0284234P.		
PR	19-APR-2001; 2001US-0285325P.		
PR	20-APR-2001; 2001US-0285381P.		
PR	20-APR-2001; 2001US-0285609P.		
PR	23-APR-2001; 2001US-0285748P.		
PR	23-APR-2001; 2001US-0285690P.		
PR	24-APR-2001; 2001US-0286068P.		
PR	25-APR-2001; 2001US-0286292P.		
PR	27-APR-2001; 2001US-0287213P.		
PR	02-MAY-2001; 2001US-0286257P.		
PR	29-MAY-2001; 2001US-0294164P.		
PR	30-MAY-2001; 2001US-0294484P.		

PR 18-JUN-2001; 2001US-0298952P.
 PR 19-JUN-2001; 2001US-0299237P.
 PR 19-JUN-2001; 2001US-0299276P.
 PR 12-SEP-2001; 2001US-0318750P.
 PR 25-SEP-2001; 2001US-0324800P.
 PR 25-SEP-2001; 2001US-0324802P.
 PR 17-OCT-2001; 2001US-0325684P.
 PR 14-NOV-2001; 2001US-0332131P.
 PR 14-NOV-2001; 2001US-0332240P.
 PR 14-NOV-2001; 2001US-0332779P.
 PR 21-NOV-2001; 2001US-0332115P.
 PR 04-DEC-2001; 2001US-0337621P.
 PR 03-JAN-2002; 2002US-0345783P.
 PR 16-JAN-2002; 2002US-0350251P.
 PR 02-APR-2002; 2002US-00114270.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Guo X, Kekuda R, Miller CE, Malysankar UM, Spylek KA;
 PI Paturajan M, Liu X, Gusev VY, Li L, Vernet CM, Zerhusen BD;
 PI Gorman L, Shenoy SC, Pena CB, Smithson G, Burgess CE, Gerlach V;
 PI Padigaru M, Shinkets RA, Gangoli BA, Taupier RJ, Caeman SJ, Ji W;
 PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ;
 PI Macdougall JR, Rothenberg MB, Mazur A, Millet I, Peyman JA;
 PI Ellerman K;
 XX
 DR WPI, 2003-046858/04.
 DR P-PSDB; ABUS4639.
 XX
 PT New isolated NOVX polypeptide useful for treating atherosclerosis,
 PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
 PT neurodegenerative disorders, Alzheimer's disease and cancer.
 XX
 PS Claim 17; Page 303; 666pp; English.
 XX
 CC The invention relates to human polypeptides, termed NOVX, and the
 CC polynucleotides encoding them. The polypeptides and polynucleotides are
 CC useful for diagnosing disease, and screening for potential therapeutic
 CC agents. The sequences are useful for treating metabolic disorders,
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
 CC stenosis, atrial septal defect (ASD), atriocentric stenosis, ventricular
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
 CC septal defect (VSD), valve diseases, tubercous sclerosis, scleroderma,
 CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX
 CC polynucleotides of the invention
 XX
 SO Sequence 1746 BP; 402 A; 457 C; 459 G; 428 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.93e-236 Length: 1746
 Score: 2597.50 Matches: 497
 Percent Similarity: 95.07% Conservative: 4
 Best Local Similarity: 94.31% Mismatches: 17
 Query Match: 84.36% Indels: 9
 DB: Gaps: 3
 US-10-023-515-2 (1-581) x ABX72267 (1-1746)
 QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPileGlnGlyLeuGln 50
 Db GGGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGGCAAGCAA 138
 QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIa 70
 Db GTCACTGCTGGAGAGCCCTGCTGTGAACGCTTCTCTGAGAGTCCCTTCTCTCTCT 198
 QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu 90
 Db CCCCCGCTGGGATCCCTCGATTTCAGAACCCGAGCCTGATGCTCCCTGGGATTAACCTTG 258

QY 91 ArgGluAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTyrPheLeuLeuAsp 110
 Db CGAGAGCCACCTCTCTACCTTATTTGGCTCCAGAACTCAGAGTGGCTGCTTAAAT 318
 QY 111 GlnHisMetLeuLeuValHisTyrProIysPheGlyValSerGluAspCysLeuTyrLeu 130
 Db CAACATATGCTCAAGTGCATTACCCGAATTCGGAGTGTGAGAAAGATGCTCTTACCTG 378
 QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerIysLeuProValLeuValTyrPhe 150
 Db AACATCTATGGCCCTGCCAGCCGATACAGGCTCCAGGCTCCGCCGTGGTGGTTC 438
 QY 151 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa 170
 Db CCAAGAGTGCCTTCAAGACTGCTCAGCTCCATCTTGAATGGTCCGCCCTGCTGCC 498
 QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
 Db TATAGAGACGTGCTGTGTGTGCTGCTCAGTACCGGCTAGGAATTTGGTTCTTCAAC 558
 QY 191 ThrTyrAspGlnHisAlaIaProGlyAsnTyrPheValAspGlnValAlaAlaLeuSer 210
 Db ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAGAGAACAGAGTGGCTGCTGTC 618
 QY 211 TrpValGlnIysAsnIleGluPhePheGlyValAspProSerSerValThrIlePheGly 230
 Db TGGGTCCAGAGAACATCATGAGTCTTCCGTGGGAACTCCAGCTCTGTACCATCTTGGC 678
 QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu 250
 Db GAGTCCGGGGAGGCATTAAGTGTCTTACTGTCTTACTGTCTCCATGCGCCAAAGCTTA 738
 QY 251 PheHisIleValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAsp 270
 Db TTCCACAAAGCATCATGAGAGTGGGTGGCCATCATCTTCACTTGAGAGCCCATAT 798
 QY 271 TyrGluIysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAsp 290
 Db TATGAGAGAGTGAAGACCTCAGGTGTGTGCATTTCTGTGTGAACATGCGTCAAC 858
 QY 291 SerGluAlaLeuLeuArgCysLeuArgThrIysProSerIysGluLeuLeuThrIysSer 310
 Db TCTGAGGCGCTCGTGAAGTGTGCTGAGGACAAACCTCCAGAGAGCTGCTGACCTGAGC 918
 QY 311 GlnIysThrIysSerPheThrArgValValAspGlyValaPhePheProAsnGluProLeu 330
 Db CAGAAACAAAGTCTTTCACCTGAGTGTGATGTGCTTCTTCTTCTTCAATGAGGCTCTTA 978
 QY 331 AspLeuLeuSerGlnIysValAlaPheValAlaIleProSerIleIleGlyValAsnAsnHis 350
 Db GATCTATGTCTCAGAAAGCATTAATAACAAATCTTCTCATCATGAGTCAATAACAC 1038
 QY 351 GluCysGlyPheLeuLeuProMetIysGlu-----AlaProGluIleLeuSerGlySer 368
 Db GAGTGTGCTTCTCTGCTGCTGCTATGTAAGATTCGTGCTGCTCCATCTGCACTCCCTCA 1098
 QY 369 AsnIysSerIleValaLeu-----HisLeuIleGlnAsnIleLeuHisIle 383
 Db AACCGTATGCAAGCTTGGCTTCAACAGCTGGGATTTCCACAGAAAG-----CATATC 1152
 QY 384 ProProGlnTyrIleHisIleValAlaAsnGluTyrPheHisAspIysHisSerIleThr 403
 Db CCGCTCTGATTTTGGACCTTGTGCTTATATAATCTTCATGACAAAGACCTCCCTGACT 1212
 QY 404 GluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValValProAlaLeu 423
 Db GAATTCGAGACAGTCTTCTGAGCTGCTTGAATGTGTTCTTGTGGTCCCTGCACTG 1272
 QY 424 IleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHis 443
 Db ATCAAGCTCGATATATCACAGAGATGCTGCTGACCTGTCTTATGAGATTTGCGGAC 1332

Qy 444 ArgProGlnCyPheGluAspThrIleProAlaPheValIleAlaSerHisAlaAspGlu 463
Db 1333 CGGCTTCAGTCTTTGAAAGACACGAGCCAGCTTTTGCAGAGCGACGCTGATGAA 1392
Qy 464 ValAlaPheValPheGlyGlyAlaPheLeuYsgIAspIleValMetPheGluGlyAla 483
Db 1393 GTCCGCTTTGTGTTCCGCTGTGCTCTCTGAGAGGGGAGCATTTTATGTTGAGAGAGCC 1452
Qy 484 ThrGluGluGluValLeuLeuSerArgIleMetIleTyrTrpAlaIlePheAlaArg 503
Db 1453 ACGAG 1512
Qy 504 ThrGluAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeuThrGluGln 523
Db 1513 ACCGGGATCTTAATGGGAGACGCTGTCTGTGGCCAGCTTATATCTGACTGAGCAG 1572
Qy 524 TyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuYsgIuProArgValAsp 543
Db 1573 TACCTCAGCTGAGCTTGAACATGAGCCTCGACAGAGACTCAAGAGACCGCGAGAGAT 1632
Qy 544 PheTrpThrSerThrIlePro 550
Db 1633 GTGTGGGTGACGGGATTCCT 1653

RESULT 10
ADRI9681
ID ADRI9681 standard; DNA; 1857 BP.
XX
AC ADRI9681;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human drug metabolizing enzyme (DME)-2 gene sequence.
XX
KW drug metabolizing enzyme; DME; cytostatic; immunosuppressive;
KW antiinflammatory; endocrine; ophthalmological; gastrointestinal;
KW hepatocarcinoma; cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; endocrine disorder; eye disorder;
KW gastrointestinal disorder; liver disorder; metabolic disorder; gene; ds;
KW human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 1..1857
FT /tag=a
FT /product="Human drug metabolizing enzyme (DME) 2"

WO200226988-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US030662.
XX
PR 29-SEP-2000; 2000US-0236947P.
PR 06-OCT-2000; 2000US-0238864P.
PR 20-OCT-2000; 2000US-0242323P.
PR 09-NOV-2000; 2000US-0247581P.
PR 16-NOV-2000; 2000US-0249519P.
PR 22-NOV-2000; 2000US-0252834P.
PR 30-NOV-2000; 2000US-0250567P.
XX
XX (INCYTE GENOMICS INC.
XX
PI Azimzai Y, Baughin MR, Borowsky MT, Ding L, Duggan BM,
PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA,
PI Lal P, Lee EA, Lu DM, Nguyen DB, Arvizu C, Policky JI, Rankumar J,
PI Rung HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK,
PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
XX
XX WPI; 2002-362498/39.
XX
XX P-PSDB; ADRI9663.

PT Human drug metabolizing enzymes, useful in the diagnosis and treatment of
PT disorders associated with aberrant (DME) activity, e.g., cancer and
PT autoimmune disorders.
XX
XX
PS Claim 12; SEQ ID NO 20; 142pp; English.
XX
XX This invention relates to novel drug metabolizing enzymes (DME) and the
CC nucleotide sequences which encode them. The invention may be useful for
CC the development of compounds with a cytostatic, immunosuppressive,
CC antiinflammatory, endocrine, ophthalmological, gastrointestinal or
CC hepatotropic activity acting as an agonist or antagonist of drug
CC metabolizing enzyme activity. The invention may be used in the diagnosis
CC and treatment of disorders associated with decreased or increased
CC expression or activity of drug metabolizing enzymes. Such disorders
CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
CC endocrine, eye, gastrointestinal (including liver disorders) and
CC metabolic disorders. The present sequence is that of a gene which encodes
CC a human drug metabolizing enzyme (DME) of the invention. Note: This
CC sequence did not form part of the printed specification but was obtained
CC in electronic format from EPO.
XX
SQ Sequence 1857 BP; 428 A; 487 C; 496 G; 446 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.52e-235 Length: 1857
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 84.07% Indels: 23
DB: 7 Gaps: 1

US-10-023-515-2 (1-581) x ADRI9681 (1-1857)

Qy 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrIleGlnGlyGln 50
Db 274 GGGCTTCTCTGAGAGGCGACAGAGAGACACAGGCTGGATGATTCAGGAGAGCA 333
Qy 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
Db 334 GTACCTGTGTGGAGAGCCCTGTGCTGTGAACGTGTCTCGAGTCCCTTCTGTCT 393
Qy 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 90
Db 394 CCCCGCTGGAGTCCCTGTGATTTAGAACCCGAGCTGCAATCCCTGGATTAATCTG 453
Qy 91 ArgGluAlaThrSerTyrProAsnLeuYsgIAsnSerGluTrpLeuLeuAsp 110
Db 454 CGAGAGGCCACCTCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGTATGAT 513
Qy 111 GlnHisMetLeuValHisTyrTrpProYsgPheGlyValSerGluAspCysLeuTyrLeu 130
Db 514 CAACATATCTCAAGGTGATTAATCCGAAATTCGAGGTGAGAGACTGCTTACCTG 573
Qy 131 AsnIleTyrAlaProAlaHisIleAlaPheThrGlySerIleYsgLeuProValIleValTrpPhe 150
Db 574 AACATCTATGCGCCCTGCGCACGCGATGACAGCTCAAGCTCCCGCTTGTGTGCTTC 633
Qy 151 ProGlyValAlaPhePheThrGlySerIleAsnIlePheAspGlySerAlaLeuAla 170
Db 634 CCGAGAGTGTCTTCAAGATGAGCTGCTCACTTCCTTGAATGGGTGCTGCTGCTGCC 693
Qy 171 TyrGluAspValLeuValAlaValAlaGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
Db 694 TATGAGAGAGT 753
Qy 191 ThrTrpAspGlnHisIleAlaProGlyAsnTrpAlaPheYsgAspGlnValAlaLeuSer 210
Db 754 ACATGGAGTCAAGATGCTCCCGGAGACTGGGCTTCAAGAGACAGGTGCTGTCTCC 813
Qy 211 TrpValGlnYsgAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
Db 814 TGGGTCCAGAGAGAGACATGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 873

```

Oy 231 GUSerAlaGlyAlaIleSerValSerLeuIleLeuSerProMetAlaValGlyLeu 250
Db 874 GAGTCCGGGGAGCCATAAGTCTTCTAGCTTATAGTCTCCCAATGGCCAAAGCTTA 933
Oy 251 PheHisValAlaIleMetGlySerGlyValAlaIleIleProTyrLeuGlnAlaHis 270
Db 934 TTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGCCCATGAT 993
Oy 271 TyrGlySerSerGlyuAspLeuGlnValAlaHisIlePheCysGlyValAspAspAspSer 290
Db 994 TATGAGAAAGAGAGAGACCTGACGGTGGTGGACATTTCTGTGTAAACATGCGTCAAG 1053
Oy 291 SerGluAlaLeuLeuArgCysLeuArgTyrHisProSerGlyLeuLeuThrLeuSer 310
Db 1054 TCTGAGGCCCTCTCTAGAGTGGCTGAGACAAACCTCCAAAGAGACTGCTGACCTCAGC 1113
Oy 311 GlnTyrThrLysSerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProLeu 330
Db 1114 CAGAAAACAAAGTCTTTCACCTGAGTGGTGTGATGGTCTTCTTCTTAATGAGCTCTTA 1173
Oy 331 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAspAspHis 350
Db 1174 GATCATATTGTCACAAAGCATTTAAAGCAATTCCTTCATCATGAGTCAATACAC 1233
Oy 351 GluCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLys 370
Db 1234 GAGTGTGGCTTCTCTCTGCTTATG----- 1257
Oy 371 SerLeuAlaLeuHisIleLeuIleGlnAsnIleLeuHisIleProProGlnIleTyrLeuHisIle 390
Db 1258 -----CACATCCCGCCCTCAGATATTGACCTT 1284
Oy 391 ValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGlnIleLysAspSerLeuLeu 410
Db 1285 GTGGGTATGATATCTTCATGACAAAGACCTCCGACTGGAATCCGAGACAGTCTTCTG 1344
Oy 411 AspLeuLeuGlyAspValPhePheValAlaProAlaLeuIleThrAlaArgTyrHisArg 430
Db 1345 GACTTGTCTGAGAGTGTCTTCTTGTGTGCTCTGACATGATACAGCTCGATATACAGA 1404
Oy 431 AspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAsp 450
Db 1405 GATGCTGGTGGACCTGTCTACTTATAGAGTTTGGGACCGGCTCAGTGTCTTGAAGAC 1464
Oy 451 ThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGly 470
Db 1465 ACGAAGCCAGCTTTTGTCAAAGCCAGCACGCTGATGAAGTCCGCTTGTGTCGTGGT 1524
Oy 471 AlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuLeu 490
Db 1525 GCCTTCTCGAAGGGGACATGTGTATGTTGCAAGAGCCAGAGAGAGAGAGAGTACTG 1584
Oy 491 SerArgLysMetMetLysTyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
Db 1585 AGCCGGAAGATATGAATATACCTGGCTACCTTGTCTCGAACCGGAAATCTTAATGGAAAC 1644
Oy 511 AspLeuSerLeuTyrProAlaTyrAsnLeuThrGluGlnIleTyrLeuGlnLeuAspLeuAsn 530
Db 1645 GACCTGTCTGTGTGGCCAGCTTATATGACTGAGAGTACTCCTCAGCTGAGACTTGAAC 1704
Oy 531 MetSerLeuGlyGlnArgLeuLysGluProArgValAspPheThrPheSerThrIlePro 550
Db 1705 ATGAGCCCTCGACAGAGACTCAAGAAACCGCGAGAGAGATGTGTGGTGAAGGGGTATCT 1764

```

```

XX cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
XX
XX
XX Felis catus.
XX
XX JP2003250575-A.
XX
XX
XX 09-SEP-2003.
XX
XX 04-MAR-2002; 2002JP-00057908.
XX
XX 04-MAR-2002; 2002JP-00057908.
XX
XX 04-MAR-2002; 2002JP-00057908.
XX
XX (TOHO-) TOHOKU TECHNOARCH KK.
XX
XX WPI, 2004-002277/01.
XX
XX P-PSDB; ADF50146.
XX
XX Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
XX
XX Claim 6; SEQ ID NO 3; 33pp; Japanese.
XX
XX
XX The present sequence encodes a cat cauxin protein (I) or its salt, which
XX is cat kidney disease marker. Also described: (1) a partial peptide (II)
XX of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
XX comprising (III); (4) a transformed host (V) comprising (III) or (IV);
XX (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
XX couples specifically with (I) or (II); (7) diagnosing cat kidney disease
XX which involves measuring (I) quantitatively, and where reduction of
XX amount of (I) indicates presence of the disease; (8) a cat kidney disease
XX diagnostic agent comprising (I) labelling agent, a reagent which measures
XX the biological activity of urinary (I) or (VI); and (9) a cauxin
XX detection kit which measures cauxin in a test sample. (I) is useful as a
XX cat kidney disease marker and (VI) is useful for diagnosing cat kidney
XX disease. (I) enables detection of cat kidney disease simply and
XX correctly. (I) provides an early marker for the disease, and replaces
XX complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
XX blood testing.
XX
XX
XX SQ Sequence 2145 BP; 444 A; 581 C; 550 G; 564 T; 0 U; 6 Other;
XX
XX
XX Alignment Scores:
XX
XX Pred. No.: 5,74e-181 Length: 2145
XX Score: 2015.00 Matches: 388
XX Percent Similarity: 80.87% Conservative: 60
XX Best Local Similarity: 70.04% Mismatches: 102
XX Query Match: 65.44% Indels: 4
XX DB: Gaps: 2
XX
XX
XX US-10-023-515-2 (1-581) x ADF50146 (1-2145)
XX
XX 31 GlyProSerAlaGlyAlaGlyProGlnArgAsnThrArgLeuGlyTyrIleGlnIlyLysGln 50
XX 228 GGGCCAGGTGGTGGAGTGCACAGTGAAGAGACACAGGCTGGAGTGGTCCGGGGAAAGCAA 297
Oy 51 ValThrValLeuGlySerProValAsnValPheLeuGlyValProPheAlaIle 70
Db 228 ACCATGTACTGGGAAGACCGGTGCTTGAACATGTTCTCCGGGATCCCTATATCTGCA 357
Oy 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu 90
Db 358 CTTCTCTAGAGCCCTCGATTTTAAGCAACCAAGCTGCTGCTCCCGGAAATGACTTC 417
Oy 91 ArgGluAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTyrLeuLeuLeuAsp 110
Db 418 CGAATGACACATCTTACCTTAATATGCTTCCAGACTTAAGAGTGGCTGCTCTAT 477
Oy 111 GlnHisMetLeuLysValHisIleTyrProLysPheGlyValSerGluAspCysLeuLysTyr 130
Db 478 CAACACGTTCTCAAGTGGTGTACCCCAATTTGAAGGCTCGGAAGACTGCTGTACCTT 537
Oy 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTyrPhe 150

```



```

Db      538 AACATCTATGCGCCAGCCCATGCGGACATAGCTCCACACCTCCCTGATAGTGTGCTTC 597
Qy      151 ProGlyGlyAlaPheLeuTherGlySerAlaSerIlePheAspGlySerAlaAla 170
Db      598 CCCGGGGGGGCTTCTGAAGATGGGCTCAGCTTCCTCTCGATGGGTCCGCTTGGCTGCC 657
Qy      171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
Db      658 TACGAGGACGCTGCTGATGCTGATCCAGTACCGGCTAGGAAATATGTTTGTTCGAC 717
Qy      191 ThrTPAPAPGlnAlaAlaProGlyAsnTPAlaPheLeuAspGlnValAlaAlaLeuSer 210
Db      718 ACAGGGGATGAGCAGATGCCCCGGGGGAACTGGGCTTGGTCGACAGGCTGGCTGCCATAC 777
Qy      211 TrpValGlnIleValAsnIleGluPhePheGlyGlyValAspProSerSerValThrIlePheGly 230
Db      778 TGGGTCCGGGACAAACATGAGATCTTGGGTGGTACCCAGCTCCGCTCCGACCATCTTGGGA 837
Qy      231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIleGlyLeu 250
Db      838 GAGTCACCGGGAGCCATCAGTGTTCAGCTCATCTGTCCCCCATAGCCATAGGCTTA 897
Qy      251 PheIleTyrAlaIleMetGluSerGlyValAlaIleIleProTyrIleu---GluAlaHis 269
Db      898 TTCACAAAGCCCATCAGAGAGAGGGGGTGGCCATCTGCTTACTGATGAGACCCCT 957
Qy      270 AspTyrGluIleSerGluAspLeuGlnValAlaAlaIlePheCysGlyAsnAsnIleSer 289
Db      958 GGTGATGAGGAAAGATGTTGACAGTGTCTTCCCGATCTGTGTGTTCCCATGGGTCT 1017
Qy      290 AspSerGluAlaLeuLeuArgCysLeuIleGlnTyrLysProSerIleGlyLeuLeuThrLeu 309
Db      1018 GACTCTGCTGCTCCCTGTCGAGCTGAGAGGCAAAACCTCCGAGAGATTGATGACATC 1077
Qy      310 SerGlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnIlePro 329
Db      1078 AGCAAGAAATCACTGTTTCCATTCAGATGATGATGATCTTTTCTTCCGATAGAGCT 1137
Qy      330 LeuAspLeuLeuSerGlnLysAlaPheLeuValaIleProSerIleIleGlyValAsnAsn 349
Db      1138 GTAGCCCTATTGACTCAAAACATTTAATTCAGTCTCTTATCATCGAGTCAATAC 1197
Qy      350 HisGluCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsn 369
Db      1198 CACGAGTGTGCTCTCTCTG---TCCACGAGTGTCTGAGATCTCTGGGGGCTCCAAAC 1254
Qy      370 LysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHis 389
Db      1255 AGGTCTGTGGCCCTTACTTACTAGTACACAGTTCCTGAAATATTTCCACCCAGTATTGTCAC 1314
Qy      390 LeuValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeu 409
Db      1315 CTGTGGCTGATCACTTCTTACAAACAGACTCCCGGTGAAATACAGAGATGATTTT 1374
Qy      410 LeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHis 429
Db      1375 CTGACCTTGCTGAGATGTGCTCTTGTGCTCCCTGGGGGTGTCACAGCTCATATCAT 1434
Qy      430 ArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnIleCysPheGlu 449
Db      1435 AGAATAGCTGTGACACTGCTTCTTATGAGTTTAAACACCCGCCCTGCTCTTAAC 1494
Qy      450 AspThrLysProAlaPheValLysValaAspHisAlaAspGluValArgPheValPheGly 469
Db      1495 GACACGAGGCGCAGCTTCTGAAAGCGGATCAGCTGATGAATCCGCTTCGCTTGGGA 1554
Qy      470 GlyAlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluValLeu 489
Db      1555 GGTGCTTCTTGAAGGCGACATGTCATGTTGAAAGGAGCCACGAGAGAGGAAATGG 1614
Qy      490 LeuSerArgLysMetLeuTyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGly 509

```

```

Db      1615 CTGAGCAGGAAAGTATGATGACTGCGCAACTTTGCTCGGACCGGGGACCTTAACGGG 1674
Qy      510 AsnAspLeuSerLeuThrProAlaTyrAsnLeuThrGlnGlnTyrLeuGlnLeuAspLeu 529
Db      1675 GAAGGTGTCTCTGTGGCGCACCTTACCCAGACGAGCAGATCTGAACCTGATTTG 1734
Qy      530 AsnMetSerLeuGlyGlnArgLeuLysGluProArgValaAspPheThrIle 549
Db      1735 AGTGTAGCGCTGTGACAGAACTGAAGAGCAAGAGGTGTGATGTAATACAAAT 1794
Qy      550 --ProLeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPhe 569
Db      1795 GTCCCTATATACCCCACTCCAGGGGCTTCCCAAGTCTCTTNNCCCTTACTCTCC 1854
Qy      569 eUSeRLeuGlnProPhePhePheCysAlaPro 581
Db      1855 TTCCTTGTCTCCCGCTTGTCTTNTTCTGTCTCA 1892

RESULT 12
ADP50144
ID   ADP50144 standard; cDNA; 1629 BP.
XX
AC   ADP50144;
XX
DT   12-FEB-2004 (first entry)
XX
DE   Cat cauxin encoding cDNA SEQ ID NO:1.
XX
KW   cat; cauxin; cat kidney disease marker; kidney disease; gene; aa.
XX
OS   Felis catus.
XX
PN   JP2003250575-A.
XX
PD   09-SEP-2003.
XX
PE   04-MAR-2002; 2002JP-00057908.
XX
PR   04-MAR-2002; 2002JP-00057908.
XX
PA   (TOHO-) TOHOKU TECHNORACH KK.
XX
DR   WPI; 2004-002277/01.
XX
PT   P-PSDB; ADP50145.
XX
PT   Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX   and for diagnosing cat kidney disease.
XX
PS   Claim 6; SEQ ID NO 1; 33p; Japanese.
XX
CC   The present sequence encodes a cat cauxin protein (I) or its salt, which
CC   is cat kidney disease marker. Also described: (I) a partial peptide (II)
CC   of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
CC   comprising (III); (4) a transformed host (V) comprising (III) or (IV);
CC   (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
CC   couples specifically with (I) or (II); (7) diagnosing cat kidney disease
CC   which involves measuring (I) quantitatively, and where reduction of
CC   amount of (I) indicates presence of the disease; and where reduction of
CC   diagnostic agent comprising (I) labelling agent; (8) a cat kidney disease
CC   detection kit which measures cauxin (I) or (VI); and (9) a cauxin
CC   cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC   disease. (I) enables detection of cat kidney disease simply and
CC   correctly. (I) provides an early marker for the disease, and replaces
CC   complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC   blood testing.
XX
SQ   Sequence 1629 BP; 345 A; 438 C; 441 G; 405 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,73e-178 Length: 1629
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60

```


Best Local Similarity: 71.73% Mismatches: 85
 Query Match: 64.47% Indels: 2
 DB: 12 Gaps: 2
 US-10-023-515-2 (1-581) x ADF50144 (1-1629)

31 GlyProSerAlaGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnIleGlySerGln 50
 64 GGGCCAGCTGCTGATGACACATGAGAGACACAGGCTGGATGGTCCGGGGAGAACAA 123
 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
 124 ACCACTGATCTGGAGACACCTGCTGGACATGTTCTCGGGATCCCATATGTGCA 183
 71 ProPheLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeu 90
 184 CCTCTCTAGGGCCCTCGATTTAAGACCAAGACCTGCTGCGGGAGATGACTTC 243
 91 ArgGlnAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTrpLeuLeuAsp 110
 244 CGAATGCGACATCTTACCTTAATATATGCTTCCAGAGATTAGAGTGGCTGCTCTAT 303
 111 GlnHisMetLeuLeuValHisTyrProIysPheGlyValSerGluAspCysLeuTyrLeu 130
 304 CAACAGCTTCTCAAGTGGCTTACCCCAATTGGAAGCTCCGAAGACTGCTGATCCTT 363
 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLeuValLeuValTrpPhe 150
 364 AACATCTATGCGCCAGCCCATGCGACATGCTCCATCCCTGTCATGATGTGTTC 423
 151 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
 424 CCGGGGGTGGCTTCAAGATGGGCTCAAGTCTCTCTCTGATGGGTCCGCTTGGCTGCC 483
 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
 484 TACGAGAGCGTGTGCTGATGCTACCCAGTACCGGCTGAGAAATTTGGTTTTCGAC 543
 191 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLeuValAsnGlnValAlaLeuSer 210
 544 ACAGGGAGTGAAGCATGCCGGGGAACTGGGCTTGTGAGACAGATGGCTCCCTCAAC 603
 211 TrpValGlnIysAsnIleGluPhePheGlyIysAspProSerSerValThrIlePheGly 230
 604 TGGGTCCGGGACACATGATGATTTCTTCGGTGTGACCCACGCTCCGTACCATTTTGA 663
 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu 250
 664 GAGTACGCGGAGCATCATGTTTCCAGCTCATTTCTGCCCATAGCCAAATGGCTTA 723
 251 PheHisValAlaIleMetGluSerGlyValAlaIleIleProTyrLeu---GluAlaHis 269
 724 TTCACAAAGCCATATGAGAGTGGGGGCTATCTCTTCTGATGAGACCCCT 783
 270 AspTyrGluIysSerGluAspLeuGlnValAlaHisPheCysGlyAsnAsnAlaSer 289
 784 GGTATGAGAGGAAGAAGATTGCGATGCTTGGCCGATCTGTGGTTGCCATGGGTCT 843
 290 AspSerGlnAlaLeuLeuArgCysLeuArgThrIysProSerIysGluLeuLeuThrLeu 309
 844 GACTCTGCTGCTCTGCTGAGTGGCTGAGGGGCAAACTCCGAGAGATGTATGACATC 903
 310 SerGlnIysThrIysSerPheThrArgValValAspGlyAlaPhePheProAsnGluPro 329
 904 AGCAAGAACTCACTTTTCCATTCAGATGATGATCTTTCTTCTTCATGAGGCT 963
 330 LeuAspLeuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAsnAsn 349
 964 GTAGCCTATTGATCAAAAGCATTTAATCATGTTCTTCTATCATCGAGATCAATAC 1023
 350 HisGluCysGlyPheLeuLeuProMetIysGlnAlaProGluIleLeuSerGlySerAsn 369
 1024 CACGAGTGTGCTTCTCTCTG---TCCACGAGTTTTCGATCATCTCGGGGCTTCAAC 1080

370 LysSerLeuAlaLeuHisIleuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHis 389
 1081 ASGTCTTGGCCCTTACTTACTTACTGACACAGCTTCTGAAATTTCCACCAAGATTTTGAC 1140
 390 LeuValAlaAsnGluTyrPheHisAspIysHisSerLeuThrGluIleArgAspSerLeu 409
 1141 CTTGGGCTGATCATTTACTTCTACAAACAGACTCCCGCGTTGAAATATAGATATATTTT 1200
 410 LeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHis 429
 1201 CTGACTTGGCTTGAGATGTGCTTTTGTGTCCCTGGGGTGTGTACAGCTGATATCAT 1260
 430 ArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGlu 449
 1261 AAGATGCTGTGACACTGTCTTCTATGAGTTTCAACACCCGCCCATGCTTAAC 1320
 450 AspThrIysProAlaPheValIysAlaAspHisAlaAspGluValArgPheValPheGly 469
 1321 GACACAGGCCAGCTTTGCTGTAAGCCATCATCTGATGAATCCGCTTGTGGA 1380
 470 GlyAlaPheLeuIysGlyAspIleValMetPheGluGlyAlaThrGluGluIysLeu 489
 1381 GGTGCTTCTCTGAAAGGCGACATGTCTATGTTCGAAAGGCGACCGAGAGAGAAATG 1440
 490 LeuSerArgIysMetMetIysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGly 509
 1441 CTGACAGAGAAATATATGATGATCTGGGCAACTTGTGCTGAGCGGGACCTTAACGGG 1500
 510 AsnAspLeuSerLeuTrpProAlaTyrAsnLeuThrGlnIysGlnIleuAspLeu 529
 1501 GAAGGTGTGCTCTGTGGCGACCTTACACCCAGAGCGAGCATTACTGAAGCTGATTTG 1560
 530 AsnMetSerLeuGlyGlnArgLeuIysGluProArgValAspPheTrpThrIle 549
 1561 AGTGTGAGCGTGGGACAGAACTGAAGGAGCAAGATGTAATTCATTCAT 1620

RESULT 13
 AB086171 standard; DNA; 1071 BP.
 AC AB086171;
 DT 10-SEP-2002 (first entry)
 XX XX
 DE Novel human gene. SBO ID 42.
 XX XX
 KW Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;
 KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
 KW antiinflammatory; cardiant; antilicer; virucide; antithyroid;
 KW ceradiprotective; anorectic; metabolic; vaccine; cancer; infection;
 KW wound healing disorders; atherosclerosis; Parkinson's disease;
 KW Alzheimer's disease; autoimmune disorder; haemotopoietic disorder;
 KW inflammation; neoplastic disease; nervous system disorder;
 KW cardiovascular disorders; pancreatitis; respiratory disorder;
 KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
 KW developmental abnormality; gastrointestinal ulceration; neuropathy;
 KW haematological disease; metabolic disease; sperm dysfunction;
 KW thyroid disorder; hypothyroidism; brain damage; colitis;
 KW cone photo- transduction deficiency; neurological disease; stroke;
 KW angioneurosis; ovulation disorder; spinal cord; thyroid gland; heart;
 KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
 KW growth abnormality; precocious puberty; gene; ss.
 XX XX
 OS Homo sapiens.
 XX XX
 PN MO200250105-A1.
 PD 27-JUN-2002.
 XX XX
 PF 17-DEC-2001; 2001WO-US049232.
 XX XX
 PR 19-DEC-2000; 2000US-0256710P.

PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
PI Mathensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX WPI; 2002-508784/54.
DR P-PSDB; ABP61006.
XX
PT Secreted proeline and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 2(a); Page 252; 335pp; English.
XX
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnary, antiarteriosclerotic, antiparkinsonian, nootropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, anticancer, virucide, antihypoid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder.
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records AB086130-AB086164 represent novel human cDNA's
CC of the invention
XX
SQ Sequence 1071 BP; 242 A; 286 C; 254 G; 289 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4,46e-159 Length: 1071
Score: 1780.00 Matches: 340
Best Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.81% Indels: 0
DB: 6 Gaps: 0
US-10-023-515-2 (1-581) x AB086171 (1-1071)
QY 242 I l e u s e r P r o m e t a l a y e g l y l e u p h e i s l y s a l a i l e m e t g l u s e r g l y v a l a 261
DB 49 A T A C T G C T C C A A G G C C A A A G G C T T A T T C C A A A G C A T C A T G A G A G G G G T G C C 108
QY 262 I l e l l e p r o t y r l e u g l u a l a h i a s p t y r g l u y s e r g l u a p l e u g l n v a l a 281
DB 109 A T C A T C C C T T A C C T G A G G C C C A T A T T A T A G A A G A C T G A C C T T C A G A G T G T G C A 168
QY 282 H i s p h e C y s g l y a n a a n u l a s e r a s p s e r g l u a l e u a l e u a r g C y a l e u a r T h r l y s 301
DB 169 C A T T T C T G T G A C A A T G C G T C A G A C T T A G A G C C C T G A G G T G C T T A G A G A C A A A 228
QY 302 P r o s e r l y s g l u e u l e u T h r l e u s e r g l n l y s T h r l y s e r P h e T h r A r g V a l a l a s p 321

DB 229 C C T T C A A G A G A C T G C T G A C C C T C A G C C A G A A A C A A A G C T T T C A C T G A G T G T G A T 288
QY 322 G l y l a p h e P h e P r o a n g l u P r o l e u a s p l e u s e r g l n l y s A l a P h e l y a l a i l e 341
DB 289 G G T G C T T C T T C C T A A T G A G C C C T T G A T A T T A T T G T C A G A A G C A T T T A A A G C A A T T 348
QY 342 P r o s e r l l e l l e g l y a l a s n a s n h i s e g l u C y s g l y P h e l e u l e u P r o m e t l y s g l u a l a 361
DB 349 C C T T C A T C A T C G A G A G T A A T A C C A G A G T G G C T T C C T G C T G C C A T A G A A G A G C T 408
QY 362 P r o g l u l l e u s e r g l y s e r a s n l y s e r l e u l a l e u h i s l e u l l e g l a s n l l e u 381
DB 409 C C T A G A G A T C C T C A G T G G C T C C A A C A A G C C C T T G C C C T C A T C G A T A C A A A A C A T C C T G 468
QY 382 H i s l l e p r o p r o g l n t y r l e u h i s l e u v a l a a n g l u t y r P h e h i s A s p l y s e r 401
DB 469 C A C A T C C G C C T C A G T A T T T G C A C C T T G G C T A T G A A T A C T T C C A T G A C A G A C A C C C 528
QY 402 L e u t h r g l u l l e a r g a s p s e r l e u a s p l e u a s p l e u g l y a s p v a l P h e P h e v a l a p r o 421
DB 529 C T G A C T G A A A T C G A G A C A G C T T C T G A C T T G C T T G A G A T G T G T C T T T G T G G C C C T 588
QY 422 A l a l e u l l e t h r a l a r g t y r h i a r g a s p l a c l y a l a p r o v a l t y r P h e t y r g l u p h e 441
DB 589 G C A C T G A T C A C A G C C G A T A T C A G A G A T G C T G T G C A C C T G T C T A C T T A T G A G A G T T 648
QY 442 A r g h i s a r g P r o g l n C y s P h e g l u a s p T h r l y s P r o l a p h e v a l l y s A l a s p h i s a l a 461
DB 649 C G G A C C G G C C T C A G T G C T T T G A A G A C A G A A G C A G C T T T T G C A A G C C G A C C A C G C T 708
QY 462 A s p g l u a l a r g P h e v a l P h e g l y l a l a P h e l e u y s g l y a s p l l e v a l m e t P h e g l u 481
DB 709 G A T A A G T C G C C T T T G T G T T G G T G G T G C C T T C G A A G G G G A C A T T G T A T G T T G A A 768
QY 482 G l y l a t h r g l u g l u l y l e u l e u s e r a r g l y s m e t c l y s t y r T r p a l a t h P h e 501
DB 769 G G A C C A C G G A G G A G A A G T T A C T A G C C G A A G A T G A A A T A C T G G G C T A A C C T T 828
QY 502 A l a a r g t h r g l y a n P r o a n g l y A s n a s p l e u s e r l e u T r p o a l a t y r A s n l e u T h r 521
DB 829 G C T G A A C C G G A A T C C T A A T G G A A G A C A C T G T C T G T G C C A G C T T A A T C T G A C T 888
QY 522 G l u g l n t y r l e u g l n l e u a s p l e u a s m e t s e r l e u g l n a r g l e u l y s g l u P r o a r g 541
DB 889 G A G A G A T A C C T C C A G C T G A C T T G A C A T G A G C C T C G G A C A G A C T C A A A G A A C C G G C G 948
QY 542 V a l a s p h e t r p h r s e r t h r l e p r o l e u l l e u s e r a l a s e r a s p m e t l e u h i s s e r 561
DB 949 G T G A T T T T G A C A C A C A C A C C A T C C C C T G A T C C T G C C T C C G A C A T C C T C A C A C G T 1008
QY 562 P r o l e u s e r l e u t h r P h e l e u s e r l e u g l n P r o p h e P h e p h e C y a l a p r o 581
DB 1009 C C T C T T T C T C T A A C T T C C T C T C T C C A G C C T T C T T T T T T T G T G T C C T 1068
RESULT 14
ADCS5523
ID ADCS5523 standard; cDNA; 1244 BP.
AC ADCS5523;
AC ADCS5523;
DT 18-DEC-2003 (first entry)
XX
DE Human carboxylatase 24.64 encoding sequence.
XX
KW human carboxylatase-24.64; primary hypertension; digestive ulcer;
KW nephrotic; bronchial asthma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 236..910
FT /*tag= a

PN CN1382799-A.
 XX 04-DEC-2002.
 XX 26-APR-2001; 2001CN-00112736.
 XX 26-APR-2001; 2001CN-00112736.
 XX (BIOW -) BIOWINDOW GENE DEV INC SHANGHAI.
 PA Mao Y, Xie Y;
 P1 MPI; 2003-269506/27.
 DR P-PSDB; ADC55524.
 XX polypeptide-human carboxylatase-24.64 and polynucleotide for coding it.
 XX Claim 6; SEQ ID NO 1, 31p; Chinese.
 CC The present invention relates to a polypeptide-human carboxylatase-24.64,
 CC the polynucleotide for coding it, the process for preparing the
 CC polypeptide by DNA recombination, the application of the polypeptide in
 CC treating diseases such as primary hypertension, digestive ulcer,
 CC nephrotic, bronchial asthma, tremor, etc, the antagonist of the polypeptide
 CC and its medical action, and the application of the polynucleotide are
 CC new. The present sequence represents human carboxylatase 24.64 encoding
 CC sequence.
 XX
 SQ Sequence 1244 BP; 301 A; 314 C; 271 G; 358 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.47e-140 Length: 1244
 Score: 1581.00 Matches: 301
 Percent Similarity: 99.67% Conservative: 0
 Best Local Similarity: 99.67% Mismatches: 1
 Query Match: 51.35% Indels: 0
 DB: Gaps: 0
 US-10-023-515-2 (1-581) x ADC55523 (1-1244)
 QY 280 ValAlaHisPheCysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuArgCysLeuArg 299
 DB 2 GTTGACATTTCTGTGGTAACAATGCGTCAGACTGAGAGCCCTGCTGAGGCTGAGG 61
 QY 300 ThrLysPProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 319
 DB 62 ACAAACCTCTCCAGAGAGCTGCTGACCCCTCAGCCAGAAACAAAGCTTCTCACTGAGTG 121
 QY 320 ValAaPGlyAlaPhePheProAsnGluProLeuAaPLeuLeuSerGlnLysAlaPheLys 339
 DB 122 GTTGAATGGTCTTCTTCTCTAATAGCCTCTAGATCTATTTGTCTCAGAAAGCATTTTAA 181
 QY 340 AlaLeuProSerLysLeuLeuValAsnAsnHisGluCysGlyPheLeuLeuProMetLys 359
 DB 182 GCAATTCCTTCATCATGAGTCAATACCAAGAGTGCGTTCTCTGCTGCTATGAG 241
 QY 360 GluAlaProGluLeuLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuLysLeu 379
 DB 242 GAGGCTCTCGAGATCTCTGAGTCTCAACAAAGTCCCTTCCATCTGATTAACAAC 301
 QY 380 IleLeuHisLeuProGluLeuLeuHisLeuValAlaAsnGluLysPheHisAspLys 399
 DB 302 ATCTGCAATCTCCGCTCTCAGATTGACCTTGGGCTTAAGAAATCACTTCATGACAG 361
 QY 400 HisSerLeuThrGluLysArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
 DB 362 CACTCCCTGAGTGAATCCGGGACAGTCTTCTGAGCTTGGAGAGATGTTCTTTGG 421
 QY 420 ValPProAlaLeuLysThrAlaArgTyrHisArgAspAlaGlyAlaProValLysPheTyr 439
 DB 422 GTCCCTGCACTGATATCAAGCTCGATATCAAGAGATGCTGGTGCACCTGTCTAATTTAT 481
 QY 440 GluPheArgHisArgProGluCysPheGluAspThrLysProAlaPheValLysAlaAsp 459

DB 482 GAGTTTCGGCACCGGCTCAGTCTTTGAAAGACAGAAAGCAGCTTTTGTCAAAGCCGAC 541
 QY 460 HisAlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspLysValMet 479
 DB 542 CACGCTGATGAGATCCGCTTTGTTCGGTGTGCTTCTTAAGAGGGGACATTTGTATG 601
 QY 480 PheGluGlyAlaThrGluGluLysLeuLeuSerArgLysMetMetLysTyrTrpAla 499
 DB 602 TTCGAAGAGGCCAGAGAGAGAGAGATTAATCTGAGCCGAGAGATATGAATATCTGGCT 661
 QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyrAsn 519
 DB 662 ACCTTTGCTCGAACCGGGAGATCTTAATGGAGACGACTGTCTGAGGCACTTAAAT 721
 QY 520 LeuThrGluGlnTyrLeuGlnLeuAspLeuMetSerLeuGlyGlnArgLeuLysGlu 539
 DB 722 CTGACTGAGCAGTACTCTCAGCTGACTTGAACATGAGCTTGGACAGACTCAAGAA 781
 QY 540 ProArgValAspPheThrThrSerThrLysProLeuLysLeuSerAlaSerAspMetLeu 559
 DB 782 CCGCGAGTGAATTTTGAACAGACACATCCCTGATCCGCTCTGCTCCAGCATGCTC 841
 QY 560 HisSerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCys 579
 DB 842 CACAGTCTCTTCTTCTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 901
 QY 580 AlaPro 581
 DB 902 GCTCCT 907
 RESULT 15
 AAF25258
 ID AAF25258 standard; cDNA; 1680 BP.
 AC AAF25258;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE Nucleotide sequence of an intestinal carboxylesterase (ICE).
 XX
 KW Intestinal carboxylesterase; ICE; tumour; cytotoxic T lymphocyte; CTL;
 KW cytotoxic factor; interleukin-2; interferon gamma; adenocarcinoma;
 KW tumour necrosis factor; cancer; hepatocarcinoma; colon; kidney; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 1..1740
 FT /*tag= a
 FT /note= "encodes AAB31700"
 FT CDS 428..928
 FT /*tag= b
 FT /note= "encodes AAB31703"
 XX
 PN MO200100784-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000MO-FR001791.
 XX
 PR 28-JUN-1999; 99FR-00008224.
 XX
 PA (INSR) INST ROUSSY GUSTAVE.
 XX
 PI Ronsin C, Scott V, Tiebel F;
 XX
 DR MPI; 2003-112443/12.
 XX
 DR P-PSDB; AAB31700; AAB31703.
 XX
 PT New peptides and its encoding nucleic acid derived from intestinal
 PT carboxylesterase, useful as immunostimulants for treating cancer.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 15, 2005, 12:02:07 ; Search time 953 Seconds
(without alignments)
3779.280 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079
Sequence: 1 MPOGLTSSASQWCFLLIQP.....PLSLTFLSLQLQPFFFCAP 581

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cg2n_1/USPTO/6664091/runat_14062005_133643_10569/app_query.fasta_1.775
-DB=Published Applications NA -OPMT=fastcap -SUFPIX=p2n.tmp -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=6664091 @CGN_1_1480 @runat_14062005_133643_10569
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cg2n_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2n_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2n_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2n_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2n_6/ptodata/1/pubpna/CTUS_NEW_PUB.seq:*
6: /cg2n_6/ptodata/1/pubpna/CTUS_PUBCOMB.seq:*
7: /cg2n_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2n_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2n_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2n_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2n_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
12: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
21: /cg2n_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
22: /cg2n_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cg2n_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
24: /cg2n_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cg2n_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cg2n_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	1746	13 US-10-023-515-3	Sequence 3, Appli
2	3079	100.0	1746	13 US-10-023-515-3	Sequence 3, Appli
3	3079	100.0	2158	18 US-10-023-515-1	Sequence 1, Appli
4	3079	100.0	2158	18 US-10-023-515-1	Sequence 1, Appli
5	3079	100.0	2158	19 US-10-023-515-1	Sequence 1, Appli
6	2915	94.7	1728	18 US-10-023-515-1	Sequence 20, Appli
7	2900	94.2	2232	18 US-10-023-515-1	Sequence 23, Appli
8	2877	93.4	2092	17 US-10-023-515-1	Sequence 24, Appli
9	2607.5	84.7	1746	18 US-10-023-515-1	Sequence 41, Appli
10	2606	84.6	1962	17 US-10-023-515-1	Sequence 736, App
11	2597.5	84.4	1746	18 US-10-023-515-1	Sequence 195, App
12	2588.5	84.1	1857	18 US-10-023-515-1	Sequence 20, Appli
13	2015	65.4	2145	20 US-10-023-515-1	Sequence 3, Appli
14	1985	64.5	1629	20 US-10-023-515-1	Sequence 1, Appli
15	1780	57.8	1071	18 US-10-023-515-1	Sequence 42, Appli
16	1230	39.9	2191	9 US-09-954-531-1038	Sequence 3854, Ap
17	1230	39.9	2191	9 US-09-954-531-1038	Sequence 1038, Ap
18	1230	39.9	2191	10 US-09-873-367C-828	Sequence 828, App
19	1230	39.9	2191	10 US-09-873-367C-828	Sequence 829, App
20	1230	39.9	2191	20 US-10-023-515-1	Sequence 27, Appli
21	1230	39.9	2191	21 US-10-023-515-1	Sequence 828, App
22	1230	39.9	2191	21 US-10-023-515-1	Sequence 829, App
23	1230	39.9	2191	21 US-10-023-515-1	Sequence 2105, Ap
24	1230	39.9	2484	9 US-09-925-298-271	Sequence 271, App
25	1230	39.9	2484	14 US-10-102-806-271	Sequence 271, App
26	1215.5	39.5	1902	9 US-09-917-800A-480	Sequence 480, App
27	1215.5	39.5	1902	9 US-10-368-934-40	Sequence 40, Appli
28	1215.5	39.5	1902	18 US-10-152-319A-531	Sequence 531, Appli
29	1181	38.4	2087	17 US-09-895-860-1	Sequence 1, Appli
30	1181	38.4	2087	17 US-10-377-072-1	Sequence 1, Appli
31	1179.5	38.3	2092	21 US-10-023-515-1	Sequence 4153, Ap
32	1179.5	38.3	3824	13 US-10-023-515-1	Sequence 22, Appli
33	1179.5	38.3	3824	13 US-10-023-515-1	Sequence 22, Appli
34	1179.5	38.3	3824	13 US-10-023-515-1	Sequence 22, Appli
35	1179.5	38.3	3824	13 US-10-023-515-1	Sequence 22, Appli
36	1179.5	38.3	3824	13 US-10-023-515-1	Sequence 22, Appli
37	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
38	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
39	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
40	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
41	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
42	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
43	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
44	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli
45	1179.5	38.3	3824	14 US-10-023-515-1	Sequence 22, Appli

ALIGNMENTS

RESULT 1
US-10-023-515-3
; Sequence 3, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy A. J.
; TITLE OF INVENTION: S105-Santago, Immaculada
; TITLE OF INVENTION: S105-Santago, Immaculada
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28

Pred. No. is the number of results predicted by chance to have a

; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1746
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-023-515-3

Alignment Scores:

Prod. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	3079.00	100.00%	100.00%	100.00%	1746	581	0	0	0	0

US-10-023-515-2 (1-581) x US-10-023-515-3 (1-1746)

```

QY      1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleuGlnPro 20
Db      1 ATGCCACAGGAACTTATCTTCATCTGCTTCAAGATGCTCTTTTCTGATTTCTCCAGCCC 60

QY      21 LeuLeuGlyYHIAArgGlnTrpGlyYstHrGlyProSerAlaGluGlyProGlnArgan 40
Db      61 CTGTGGGACACAGACAGTGGGGAAAACCTGGCTTCTGCTGAGGGCCACAGAGAAC 120

QY      41 ThrArgLeuGlyYTrpIleGlnGlyYsGlnValThrValLeuGlySerProValProVal 60
Db      121 ACCAGGCTGGAGTGGATTCAGGGCAAGCAAGTCACTGTCTGGAAAGCCCTGTGCTGTG 180

QY      61 AasnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
Db      181 AACGTGTTCTCGAGAGTCCCTTTGTCTGCCCGCTGGAGTCCCTGGATTTCAGAAC 240

QY      81 ProGlnProAlaSerProTrpPhePheLeuArgGlnValAlaThrSerTrpProAlaLeuGly 100
Db      241 CCGCAAGCTGCAATCGCTCTGGATTAACCTTGGAGAGCCACCTCTTACCTTAATTTGTGC 300

QY      101 LeuGlnAasnSerGlnTrpLeuLeuLeuArgGlnHISmetLeuYsValHISTrpProYs 120
Db      301 CTCGAGAACTCAGAGTGGCTGCTTGAATCAACATGCTCAAGGTGATTAACCCGAAA 360

QY      121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTyraAlaProAlaHISAlaAspThr 140
Db      361 TTCGAGAGTGCAGAACTGCTCTTACCTGAACATCTAGCGCTGCCACGCGCATACA 420

QY      141 GlySerYsLeuProValLeuValTrpPheProGlyValAlaPheYsThrGlySerAla 160
Db      421 GGCTCCAAAGCTCCCGCTTGTGTGTGTTCCAGAGAGGTGCTTCCAAAGCTGGCTCAAGC 480

QY      161 SerIlePheAspGlySerAlaLeuAlaAlaTyrgIuAspValLeuValValValGln 180
Db      481 TCCATCTTGAATGGGTCCGCCCTGGCTGCTATGAGGAGCGTGGTGTGTGTGTGTGTGT 540

QY      181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHISAlaProGlyYasnTrp 200
Db      541 TACGGGCTAGAGAAATTTTGTGTTCTTCCACACATGGAGTCAAGATGCTCCGGGAACTGG 600

QY      201 AlaPheYsAspGlnValAlaAlaLeuSerTrpValGlnYsAsnIleGluPhePheGly 220
Db      601 GCTTTCAAAGACACAGTGGCTGCTGTCTCGGTCCAGAAAGAACATGCAATTTCTGGGT 660

QY      221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer 240
Db      661 GGGGAGCCCAAGCTCTGTGACCATCTTGGGAGTCCGGGGAGGCATTAAGTGTTCAGT 720

QY      241 LeuIleLeuSerProMetAlaYsGlyLeuPheHISYsAlaAlaIleMetGluSerGlyVal 260
Db      721 CTTATACCTGCTCCCAAGGCCAAAGGCTTATTCACAAAGCAACATGAGAGAGGGGTG 780

QY      261 AlaIleIleProYsLeuGlnValAlaHISAspTyrgIuYsSerGluAspLeuGlnValVal 280
Db      781 GCCATCATCCCTTACCTGAGAGGCCCATGATTATGAGAAAGTGAAGCATCTGCAAGTGTGT 840
  
```

```

QY      281 AlaHISpeCySGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
Db      841 GCACATTTCTGTGTGTAACAATGCGTCAAGCTCTAGAGGCTCTGTCAAGTCTGAGAGCA 900

QY      301 YsProSerYsGluLeuLeuThrIleuSerGlnYsThrYsSerPheThrArgValVal 320
Db      901 AAACCTCCAAAGACCTGCTGACCTCCAGCAAGAAACAAAGTCTTTTCACTGAGTGGTT 960

QY      321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnYsValAlaPheYsAla 340
Db      961 GATGCTGCTTTCTTTCTTAATGAGCTTCAATCTATTGTTCTCAGAAAGCATTTAAAGCA 1020

QY      341 IleProSerIleIleGlyValAsnAsnHISGluCysGlyPheLeuLeuProMetYsGlu 360
Db      1021 ATTCCCTTCATCATCGAGTCAATTAACAACAGAGTGTGGCTTCTGTGCTGCTATGAAGAG 1080

QY      361 AlaProGluIleLeuSerGlySerAsnYsSerLeuAlaLeuHISLeuIleGlnAsnIle 380
Db      1081 GCTCCTGAGATCTCAAGTGGCTCCAAACAAAGTCCCTTGGCCCTCATCTGATACAAAGATC 1140

QY      381 LeuHISIleProProGlnTrpLeuHISLeuValAlaAsnGluYrPheHISAspYsHIS 400
Db      1141 CTGACACATCCCGCTCAGATATTGCACTTGTGGCTATGAATCTTCATGACAAAGCAC 1200

QY      401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
Db      1201 TCCCTGACTGAATCCAGACAGTCTTGTGACTTGTGGAGATGTGTCTTGTGTGTCTC 1260

QY      421 ProAlaLeuIleThrAlaArgYrHISArgAspAlaGlyAlaProValYrPheYsGlu 440
Db      1261 CTTGCACTGATCAACAGTTCATATACAGAAATGCTGTGCACTGTCTTACTTCAAGAG 1320

QY      441 PheArgHISArgProGlnCysPheGluAspThrYsAspProAlaPheValYsAlaAspHIS 460
Db      1321 TTTGGGACCGGCTCAGTCTTGAAGACACGAAAGCCGCTTTTGTCAAAAGCGAACAC 1380

QY      461 AlaAspGluValArgPheValPheGlyValAlaPheLeuYsGluYsAspIleValMetPhe 480
Db      1381 GCTGATAAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440

QY      481 GlnGlyValaThrGlnGlnGlnYsLeuLeuSerArgYsMetMetYsYrTrpAlaThr 500
Db      1441 GAAGGACCCACGAGAGGAGGAAAGTTACTAGAGCCGGAAGATGATGAATCTGGGCTAAC 1500

QY      501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaYrAsnLeu 520
Db      1501 TTTGCTCGAAACCGGAAATCTTAATGGGAAACGACTGTCTGTGGCCAGCTTATTAATCTG 1560

QY      521 ThrGlnIleTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuYsGluPro 540
Db      1561 ACTGAGCAATCTCCAGCTGAGCTTGAACATGAGCTTGGAGAGAGACTCAAAAGAACCG 1620

QY      541 ArgValAspPheThrPheSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHIS 560
Db      1621 CGGGTGAATTTTGAACACAGACATCCCGCTGATCTGTGCTCCGACATGCTCCAC 1680

QY      561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCyAla 580
Db      1681 AGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

QY      581 Pro 581
Db      1741 CCT 1743
  
```

RESULT 2

US-10-674-636-3
 ; Sequence 3, Application US/10674636
 ; Publication No. US20040086922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Curtis, Roy A. J.
 ; APPLICANT: Silos-Santiago, Immaculada
 ; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE

```

1  TITLE OF INVENTION:  FAMILY MEMBER AND USES THEREOF
2  FILE REFERENCE:  10448-122001
3  CURRENT APPLICATION NUMBER:  US/10/674,636
4  CURRENT FILING DATE:  2003-09-29
5  PRIOR APPLICATION NUMBER:  US/10/023,515
6  PRIOR FILING DATE:  2001-12-18
7  PRIOR APPLICATION NUMBER:  60/256,369
8  PRIOR FILING DATE:  2000-12-18
9  PRIOR APPLICATION NUMBER:  60/219,508
10 PRIOR FILING DATE:  2001-03-28
11 NUMBER OF SEQ ID NOS:  6
12 SOFTWARE:  FastSeq for Windows Version 4.0
13 SEQ ID NO 3
14 LENGTH:  1746
15 TYPE:  DNA
16 ORGANISM:  Homo sapiens
17 US-10-674-636-3

```

Alignment Scores:	
Pred. No.:	0
Score:	3079.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	18
US-10-023-515-2 (1-561)	x US-10-674-636-3 (1-1746)
	Length: 1746
	Matches: 561
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

Qy	1	MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuLeuGlnPro	20
Db	1	ATGCCACAGGAGCTTACTTCACTGCTTCAACAATGGGCTTTTCTGATCTTCCAGCCC	60
Qy	21	LeuLeuGlnYhiSaArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgAsn	40
Db	61	CTGTGTGGACACAGACAGTGGGGAAAACTGGGCCCTTCTGCTGAAGGGCCACAGAGAAC	120
Qy	41	ThrArgLeuGlnYTrpLleGlnGlyLeuGlnValThrValLeuGlySerProValProVal	60
Db	121	ACCAAGCTGGAGTGAATTCAGGACCAAGTCACTGTGTGGAAAGCCCTGTGCTGTG	180
Qy	61	AsnValPheLeuGlnYValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn	80
Db	181	AACGTGTTCTCTGGAGTGCCCTTTGTGTGCTCCCGCTGGAGTCCCTGGCATTTAGAAC	240
Qy	81	ProGlnProAlaSerProTrpAspAsnLeuArgGlnAlaThrSerTrpProAsnLeuCys	100
Db	241	CCGACGCCCTGCATCCCTCCGGGATTAATTGGCGAAGGCCACCTCTTAATTTGTGTC	300
Qy	101	LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuLeuValHisTrpProLys	120
Db	301	CTCCGAAGCTCAGATGGCTGCTCTTAAGATCAACAATGCTCAAGGTGCAATTACCCGAAA	360
Qy	121	PheGlnYValSerGlyAspCysLeuTrpLeuAsnLleTrpAlaProAlaHisAlaAspThr	140
Db	361	TTCCGAGTGTCAAGAAGACTGCCCTCACTGAACATCTAATGCCGCTGCCACGCCGATACA	420
Qy	141	GlySerTrpLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySerAla	160
Db	421	GGCTTCAAGCTCCCGCTCTGTGTGTGGTTCCACGAGGAGTCCCTTCAAGATGGCTCAGCC	480
Qy	161	SerLlePheAspGlySerAlaLeuAlaAlaTrpGlyAspValLeuValValAlaGln	180
Db	481	TCCATCTTTGATGGGTCGCCCTGTGCTGCTAATGAGACGTGTGTGTGTGTGCTGCCG	540
Qy	181	TyrArgLeuGlnYlLePheGlyPhePheThrTrpTrpAspGlnHisAlaProGlyAsnTrp	200
Db	541	TACCCGCTAAGAAATATTGGTTCTTCAACACATGGAGTACAGATCTCCGGGGAACTGG	600
Qy	201	AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnLleGlnPhePheGly	220
Db	601	GCCTTCAGAGCACAGGTGCTGCTCTGTCTGGGTCCAGAAAGAACATCGAATTTCTTCGGT	660
Qy	221	GlyAspProSerSerValThrLlePheGlyGlnSerAlaGlyAlaLleSerValSerSer	240

Dd	661	GGGAGCCCACTCTGACCACTCTTGGCGAATCCCGGAGCCATGAAGCTTCTAAGT	720
Qy	241	LeuIleuSerPrometalalyGlyLeuPhehiValysalailIemebIuSerGlyVal	260
Dd	721	CTTAATACTGTCTCCCATGCGCAAGGCTTATTCCAAAGCCATCATGAAAGTGGGGTG	780
Qy	261	AlaIleIleProIyTLeuGluIalAhiIaSPyTgIuIySSerGluIaPheGluIVal	280
Dd	781	GCATATCACTCCCTTACCTGGAGGCCCATGATTATAGAAAGATGAGACCTGACAGTGGTT	840
Qy	281	AlaIhSPheCySGIYAsnAsnAlaSerAspSerGluIalaleuIaIaGyCyLeuArgThr	300
Dd	841	GCAATTTCTGTGGTAACATATCGCTCAAGCTCTAGAGCCCTGAGGTGCTTAGAGACA	900
Qy	301	LysProSerIySGIuLeuLeuThrLeuSerGluIuIyThIySerpheThArgValVal	320
Dd	901	AAACCTTCAGAGAGAGCTGCTACCCCTGACGCCAGAAAACAAGCTTTTCACTCGAGTGGTT	960
Qy	321	AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGluIyAlaPheIyAla	340
Dd	961	GATGGTGCTTTCTTCTTCCATATGACCCCTCAATCATATTTCTCGAAAGCATTTTAAAGCA	1020
Qy	341	IleProSerIleIleGlyValAsnAsnIhISgluCySGIYPheLeuAspPomerIySGlu	360
Dd	1021	ATTCCTTCATCAATCGAGATCAATTAACACACAGTGTGGCTTCCTGCTGCTTATAGAGAG	1080
Qy	361	AlaProGluIleLeuSerGlySerAsnIySerLeuAlaLeuIhISleuIleGluAsnIle	380
Dd	1081	GCTCTGAGATCCCTCAAGTGGCTCCAAACAAGCCCTTGCCCTCATCTGATCAAAACATC	1140
Qy	381	LeuIhISIleProProGluIyTLeuIhISleuValAlaAsnGluIyTYPheIhISaSPlyhIS	400
Dd	1141	CTGCAACATCCCGGCTCTAGTATTTGACACTTGCTGCTATGATCTTCCATGACAAACAC	1200
Qy	401	SerIeuThrgIuIleArgAspSerLeuLeuAspLeuGluYAspValPhePheValVal	420
Dd	1201	TCCCTGACTGAAATCCAGACAGCTTCTGACTGCTTGGAGATGAGTCTTGTGGTCTC	1260
Qy	421	ProIalauIleThrIalArgTYhISArgAspAlaGlyAlaProValIyTYPheTYGlu	440
Dd	1261	CTTCACCTGATCAGACAGCTCATATCACAGAGATCTGTGTGACCTGTCTACTTATAGAG	1320
Qy	441	PheIhArgIhISArgProGluCySPheGluAspThIyAspProAlaPheValIyAlaAspIhIS	460
Dd	1321	TTTGGGCAACCGGCTCTCAGTGTCTTGAAGACACGAAGCCGGCTTTGTCAAAAGCCGACAC	1380
Qy	461	AlaAspGluValaArgPheValPheGlyGlyAlaPheLeuIySGIYAspIleValMetPhe	480
Dd	1381	GCTATGATGAAGTCCGCTTGTGTGGTGGTGGCTCTTCTGAAAGGGGGAATGTTATGTCTC	1440
Qy	481	GluGluIalAThrgIuGluGluIyLeuLeuSerArgIySmetIeIyTYTTPAlaThr	500
Dd	1441	GAAAGAGCCACGAGAGAGAGAGAACTTACTGAGCCGGAAGATGAAATACTGGGCTAAC	1500
Qy	501	PheAlaArgThrgIyAsnProAsnGlyAsnAspLeuSerLeuTTPProAlaIyTAsnIleu	520
Dd	1501	TTTGTCTGAAACCGGGAATCTTAATGGGAAGACCTGTCTGTGGCCAGCTTATATCTG	1560
Qy	521	ThrgIuGluIyTLeuGluIleuAspLeuAsnMetSerLeuGlyAlaGluIySGluPro	540
Dd	1561	ACTAGACAGTATCCCTCCAGCTGAGACTTGAACATGAGCTTCGAGACGAATCTCAAGAACG	1620
Qy	541	ArgValAspPheTTPThrSerThIleProLeuIleuSerAlaSerAspMetLeuIhIS	560
Dd	1621	CGGGTGAATTTTGGACACAGACCAATCCCTCTGATCTGTCTGCTCCGACATCTCCAC	1680
Qy	561	SerProLeuSerSerLeuThrPheLeuSerLeuGluIuProPhePhePheCyAla	580
Dd	1681	AGTCTCTCTTCTCTTAATCTTCTCTCTCTCTCTCAGCTTTCTTTTCTTTTGAGCT	1740
Qy	581	Pro 581	


```
Db          1741 CCT 1743
RESULT 3
US-10-023-515-1
; Sequence 1, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Siles-Santiago, Inmaculada
; TITLE OF INVENTION: 53010. A NOVEL HUMAN CARBOXYLESTERASE
; FILE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/023, 515
; PRIORITY FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-023-515-1

Alignment Scores:
Pred. No.:          0          Length:      2158
Score:             3079.00      Matches:      581
Percent Similarity: 100.00%     Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:       100.00%     Indels:      0
DB:                13          Gaps:          0

US-10-023-515-2 (1-581) x US-10-023-515-1 (1-2158)

QY      1 MetProGInGlyLeuThiSerSerAlaSerGInTPCyPhePheLeuIleuGInPro 20
Db      96 ATGCCACAGAGGACTTACCTTCATCTGCTTCACAAATGGTCTTTCTGATTCCTCAGGCC 155
QY      21 LeuLeuGlyVHIArGInTPrGlyVbThrGlyProSerAlaGluGlyProGInAArgAsn 40
Db      156 CTGTTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGGAAAC 215
QY      41 ThrArgLeuGlyTTPrlEgInGlyLySgInValThrValLeuGlySerProValProVal 60
Db      216 ACCAGGCTGGGATGGATTCAAGGGCAAGCACTGTGCTGGGAAAGCCCTGTGCTGTG 275
QY      61 AenValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuAArgPheThrAsn 80
Db      276 AACGTTCTCTGGAGTCCCTCTTGCTGCTCTCCCGCTGGGATCCCTGCGATTTCAGAAC 335
QY      81 ProGInProAlaSerProTTPAAspAsnLeuArgGluAlaThrSerTyrProAsnLeuCyS 100
Db      336 CCGGAGCTGCATCGCCCTGGGATTAACCTTGGAGAAGCACTCTTACCTTAATTTGTGC 395
QY      101 LeuGInAsnSerGlyUTPrPLeuLeuLeuAspGlnHisMetLeuLyValHisTyrProLyS 120
Db      396 CTCGAAGAACTAGAGTGGTCTCTTATGATCAACATGCTCAAGGTGCAATTACCGGAA 455
QY      121 PheGlyValSerGluAspCySLeuTyrLeuAsnIleTyrAlaProAlaHisAlaAspThr 140
Db      456 TTGCGAGTGTAGAAAGCTGCTTACTTACCTTAACATCTTAAGGCTGCGCAAGCGGATACA 515
QY      141 GlySerLeuLeuProValLeuValTTPPheProGlyValAlaPheLySThrGlySerAla 160
Db      516 GGGTCCAAAGTCCCGCTTGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGGCTACGCC 575
QY      161 SerIlePheAspGlySerAlaLeuAlaAlaTyrGluAspValLeuValValGln 180
Db      576 TCCATCTTGAATGGGTCCGCCCTGGCTGCTATGAGAGAGTGTGTTGTGTGCTCCAG 635
QY      181 TyrArgLeuGlyTlEPheGlyPhePheThrTTPAAspGlnHisAlaProGlyAsnTP 200
Db      636 TACCGCTAGGAATTTTGGTTTCTTACCAATGGATGACAGTCCCGGGAACCTGG 695
QY      201 AlaPheLyAspGInValAlaAlaLeuSerTPValGlnLySAsnIleGluPhePheGly 220
Db      696 GCTTTCAGAGACCAAGTGGTCTGTCTGTGGGTCCAGAAACATGCAATTTCTGCGT 755
QY      221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer 240
Db      756 GGGAGACCCCAAGCTGTGATCAATCTTTGGCCAGATCCCGGGAGCCATTAAGTCTTAGT 815
QY      241 LeuIleLeuSerProMetAlaLySgLyLeuPheHisLySAlaIleMetGluSerGlyVal 260
Db      816 CTTATATGCTCTCCCATATGCCAAAGGCTTAATCCAAAGCCATCATGAGAGTGGGGT 875
QY      261 AlaIleIleProTyrTLeuGluAlaHisAspTyrGluLySserGluAspLeuGInVal 280
Db      876 GCCATCATCCCTTACCTGGAGGCCCATGATTAAGAAAGTGAAGAGTGCAGAGTGCTT 935
QY      281 AlaHisPheCySgLyAsnAsnAlaSerAspSerGluAlaLeuLeuAArgCySLeuAArgThr 300
Db      936 GCACATTTCTGTGTAAACATGCTGTAGACTGTAGGCCCTGTGAGTGTCTTGAAGACA 995
QY      301 LySProSerLySgLyLeuLeuThrLeuSerGlnLybThrLySserPheThrArgVal 320
Db      996 AAACCTTCAGAGAGCTGTGACCTTCAGCCACAGAAACAAAGCTTTCATCCAGAGTGTT 1055
QY      321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLySAlaPheLySAla 340
Db      1056 GATGCTCTTTCTTTCTTAATAGCTCTCAACAAAGTCCCTGCTCCATCTGTAACAAACATC 1115
QY      341 IleProSerIleIleGlyValaAsnAsnHisGlyCySgLyPheLeuLeuProMetLySgIn 360
Db      1116 ATTCTTCCATCATCGAGTCAATTAACCAAGATGTGGCTCTGCTGCTATGAAAGAG 1175
QY      361 AlaProGluIleLeuSerGlySerAsnLySserLeuAlaLeuHisIleuIleGlnAsnIle 380
Db      1176 GCTCTGAGATCTCTCAAGTGTCTCAACAAAGTCCCTGCTCCATCTGTAACAAACATC 1235
QY      381 LeuHisIleProProGlnTyrLeuHisLeuValAlaAsnGlyTyrPheHisAspLySHis 400
Db      1236 CTGCACATCCCGCTCAGTATTGCACTTGTGGTAAAGATCTTCATACAAAGCAC 1295
QY      401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 420
Db      1296 TCCCTAAGTAATCCGAGACAGTCTTGTGACTTGTGGAGATGTCTTGTGTGTC 1355
QY      421 ProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGlu 440
Db      1356 CTGCACATGATCAAGCTGATATCAAGAGATGTGTGGGACCTGTCTACTCTTATAG 1415
QY      441 PheArgHisArgProGlnCySAspGluAspThrLySProAlaPheValAlaAspHis 460
Db      1416 TTTGGCAACCGGCTCAAGTCTTGAAGACAGAAAGCCGCTTTTGTCAAAAGCCGACAC 1475
QY      461 AlaAspGluValaArgPheValPheGlyGlyAlaPheLeuLySgLyAspIleValIlePhe 480
Db      1476 GCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1535
QY      481 GluGlyAlaThrGluGluGluGlyLeuLeuSerArgLySMeMetLySyrTTPAlaThr 500
Db      1536 GAAGAGCCACAGAGAGAGAGAGTACTGAGCCGGAAGAGATGAAATACTGGGCTAAC 1595
QY      501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTTPProAlaTyrAsnLeu 520
Db      1596 TTTGCTCAACCGGGAATCTTAATGGGAACGACTGTCTGTGTGCGACGCTTAATATCTG 1655
QY      521 ThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLyLeuLySgInPro 540
Db      1656 ACTGAGAGTACCTTCCAGCTGGACTTGAACATGAGCTTCGAGACAGAGACTCAAGAACCG 1715
```


QY 541 ArgValaerPheTrpThrSerThrIlePheLeuIleLeuSerAlaSerAspMetLeuHis 560
DB 1716 CCGGAGATTGTTGGACGACGACATCCCCCTGATCTCTGCTCCGACAGTCCGAC 1775
QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
DB 1776 AGTCTCTCTTCTTCTTAACTTCTCCTCTCTCTCCGACCTTTCTTTTCTGTGCT 1835
QY 581 Pro 581
DB 1836 CCT 1838
RESULT 4
US-10-674-636-1
; Sequence 1, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-12201
; CURRENT APPLICATION NUMBER: US/10/674, 636
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279, 508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-674-636-1
Alignment Scores:
Pred. No.: 0 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-10-023-515-2 (1-581) x US-10-674-636-1 (1-2158)
QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
DB 96 ATGCCACAGGAGCATTAATTCATCTGCTTCAACAATGCTGCTTTTCTGATTTCTCCAGCCC 155
QY 21 LeuLeuGlyHisArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgan 40
DB 156 CTGTTGGGACACAGACAGTGGGAAAACCTGGGCCCTTCTGTCGAGAGGCGCACAGAGAAC 215
QY 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
DB 216 ACCAGGCTGGAGTGAATTCAGGCGACAGTCACTGCTCTGGAGAGCCCTGCTGTG 275
QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
DB 276 AACGCTTCTCTGGAGTCCCTCTTGTCTCTCCGCCCTGGATCCCTGGAGTTTACGAC 335
QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnValThrSerTrpAsnLeuCys 100
DB 336 CCGCAGCCTGCATCGCCCTGGAGTAACTTGGCGAAGACCACTCCCTAATCCCTAATTTGTGC 395
QY 101 LeuGlnAsnSerGluTrpLeuLeuLeuAspGlnHisMetLeuLysValHisTrpProLys 120

DB 396 CTCGAGAACTCAGAGTGGCTGCTTAATGATCAACATGCTCAAGTGCATTACCGGAA 455
QY 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
DB 456 TTCGAGAGTCAAGAGACTGCTCTTAACATCTTGTGCCCTGGCCCAACCGCATACA 515
QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
DB 516 GGCTCCAGACTCCCGCTTGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGCTCAGGC 575
QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValGln 180
DB 576 TCCATCTTTGAGTGGATCCGCCCTGGCTGCTATGAGAGCTGCTGTGTGTCATCAG 635
QY 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHisAlaProGlyLeuTrp 200
DB 636 TACCGGCTAAGAAATTTGTGTTCTTCAACATGAGATCAGACAGTCCCGGGAACTGG 695
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGluPhePheGly 220
DB 696 GCCTTCAGAGACCAAGTGGCTGCTGTCTGTGGATCCAGAGAACATCGAGTTCTCGGT 755
QY 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
DB 756 GGGAGACCCAGCTCTGTACCATCTTTGGCGAGTCCGGGGAGCATTAAGTCTTACTAGT 815
QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysValIleMetGlnSerGlyVal 260
DB 816 CTTAATACGTCTCCCATGCGCCAAAGCTTATCCAAAGCATCATGAGATGGTGGG 875
QY 261 AlaIleIleProTrpLeuGlnValHisAspTrpGlyLysSerGlnAspLeuGlnValVal 280
DB 876 GCCATCATCTCTTACCTGAGGCCCATGATTTATGAAAGATGAGAGACTTGAGGTGCTT 935
QY 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
DB 936 GCACATTTCTGTGTGATACATGCGTCAAGCTTGAGGCCCTGCTGAGGTCCTGAGAC 995
QY 301 LysProSerLysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVal 320
DB 996 AAACCTCCAAAGAGACTGCTGACCTCCAGCCGAAACAAAGTCTTCACTGAGTGGTT 1055
QY 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
DB 1056 GATGATGCTTCTTTCCTTAATGAGCTCTAAGTCTAATGCTCAAGAACATTTAAAGCA 1115
QY 341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLysGln 360
DB 1116 ATTCCTTCATCATCGAGTCAATAACACAGATGCTGCTCTGCTGCTATGAAGAG 1175
QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIle 380
DB 1176 GCTCTGAGATCTCTGAGTGGCTCAACAAAGTCCCTTGGCTCATCTGATACAAAACATC 1235
QY 381 LeuHisIleProProGlnTrpIleuHisLeuValAlaAsnGluTrpPheHisAspLysHis 400
DB 1236 CTGCACATCCCGCTCAGATTTTGCACCTTGGGCTAAGTAATCTTCCATACAAAGCAC 1295
QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
DB 1296 TCCCTGATCGAATTCGAGACAGTCTTCTGACCTTGGAGAGTGTCTTGTGGTGC 1355
QY 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGlu 440
DB 1356 CCTGACATGATCAAGCTCGATATCAAGAGATGCTGGTGCACCTGTCTACTATATGAG 1415
QY 441 PheArgHisArgProGlnCysPheGlyLysPheTrpLysProAlaPheValLysAlaAspHis 460
DB 1416 TTTTCGACACCGGCTCAGGCTTTTGAAGACAGAAACCGGCTTTGTCAAAACCAACAC 1475
QY 461 AlaAspGlyValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480

Db 1476 GCTGATGAAAGTCCGCTTGTGTGTCGGTGGCTTCCTCGAAGGGGACATTGTATGTTTC 1535
Qy 481 GUGUyAlaThrGUGUyGUGUyLyuLeuSerArgLysMetClyGTYTTPAlaThr 500
Db 1536 GAAGAGACCGAGGAGAGAGAAAGTACTGAGCGGAGAAAGTGAATGAATCTGGGCTAAC 1595
Qy 501 PheAlaArgThrGlyAenProAsnGlyAsnAspLeuSerLeuTPProAlaTYAsnLeu 520
Db 1596 TTTCCTCGAACCGGGAATCTCAATAGGACACGACTGTCTGTGGCCAGCTTTATATCTG 1655
Qy 521 ThrGUGUyLyuLeuGUGUyLyuLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Db 1656 ACTGAGACGATACCTCCAGCTGAGACTTGACATGAGCTCCGACAGACACTCAAGAACCG 1715
Qy 541 ArgValaAspPheThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
Db 1716 CCGGTGATTTTGGACACACCATCCCTGATCTGTCTCTCCGACAGCTCCAC 1775
Qy 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
Db 1776 AGTCCCTTCTTCTTAACCTTCTCTCTCTCTCCAGCTTTCTTTTCTTTGTGCT 1835
Qy 581 Pro 581
Db 1836 CCT 1838

RESULT 5
US-10-757-262-45
; Sequence 45, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:

; APPLICANT: Katichefi, Venkateswarlu
; APPLICANT: Siles-Santiago, Imaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 55872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 22077, 22245, 2387, 52908, 62112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; FILE REFERENCE: MP103-007P1RMONM1M
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: CDS
; LOCATION: (96) ... (1841)
US-10-757-262-45

Alignment Scores:
Pred. No.: 0 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-023-515-2 (1-581) x US-10-757-262-45 (1-2158)

Qy 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
Db 96 ATGCCACAGGACCTTACTTCACTGCTTCCACAAATGGGCTTTTCCGATCTCCAGCCC 155
Qy 21 LeuLeuGlyLysArgGlnTrpGlyLysThrGlyProSerAlaGlnGlyProGlnArgAsn 40
Db 156 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAC 215
Qy 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
Db 216 ACCAGGCTGGGATGATTCAGGCGACAGCACTGTGTGGGAACCTGTGCTGTG 275
Qy 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
Db 276 AACGTGTCTCGGAGTCCCTTTGCTGTCTCCCGCTGGGATCCTTGCAATTTACGAC 335
Qy 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnAlaThrSerTrpProAsnLeuCys 100
Db 336 CCGAGGCTGATCGCCCTGAGTAACCTTGCAGAAAGCCACTCCATATTTGTGC 395
Qy 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuLysValHisTrpProLys 120
Db 396 CTCAGAGACTCAGATGGCTGCTCTTATGATCAACATGCTCAAGGTGCTTACCCGAAA 455
Qy 121 PheGlyValSerGlyAspCysLeuTrpLeuAsnIleTyraLeuProAlaHisAlaAspThr 140
Db 456 TTGGAGTGCACAAAGACGCTCTCACTGACATCATGAGCGCTGCCACGCCGATACA 515
Qy 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
Db 516 GGTCTCAAGTCCCTCTTGTGATGATGATCCAGAGGGGCTTCAAGACTGCTCAGCC 575
Qy 161 SerIlePheAspGlySerAlaLeuAlaAlaTyrgLysAspValLeuValValGln 180
Db 576 TCCATCTTGAATGGGTCCGCCCTGGCTGTATGAGGACGTCTGTTGTGCTCCAG 635
Qy 181 TyraArgLeuGlyLysPheGlyPhePheThrTrpAspGlnHisAlaProGlyAsnTrp 200
Db 636 TACCGGCTAGGAATATTTGGTTCTTCCACACATGGGATCAGCATGCCGGGAACTGG 695
Qy 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGlnPhePheGly 220
Db 696 GCCTTCAAGGACCAAGTGGCTGTCTGTCTGCTGCTCCAGAAACATGAGTTCTTCGT 755
Qy 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
Db 756 GGGAGCCCAAGCTGTGACCATCTTGGCGAGTCCGGGAGGCCATAAGTTCCTAGT 815
Qy 241 LeuIleLeuSerProMetAlaLysGlyLysPheHisLysValIleMetGlnSerGlyVal 260
Db 816 CTTAATCTGTCTCCATGGCCAAAGCTTATTCCAAAGCCATCTGTGAAGATGGGGGTG 875
Qy 261 AlaIleIleProTyraLeuGlnAlaHisAspTyrgLysSerGlyAspLeuGlnValVal 280
Db 876 GCCATCATCCCTTACTGAGGCGCATATATGAAAGATGAGGACCTCCAGGTGTT 935
Qy 281 AlaHisPheCysGlyValAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300

```

Db 936 GCACATTTCTGTGTAAACATGCTGACATCTGAGCCCTGAGTGCCTGAGACA 995
Qy 301 LysPProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
Db 996 AAACCTTCAGAGAGCTGTGACCTTCAGCCAGAAACAAAGCTTTTCACTGAGTGGTT 1055
Qy 321 AapGlyAlaPhePheProAngLupProLeuAapLeuLeuSerGlnLysAlaPheLysAla 340
Db 1056 GATGTGCTTTCTTTCTTAATGAGCCTAGATCTATTGCTCCAGAAAGCATTTAAAGCA 1115
Qy 341 ILeproSerLLeIleGlyValAsnAsnHisGluCysGlyPheLeuLeuProMetLysGlu 360
Db 1116 ATTCCTTCATCATGAGTCAATTAACCAAGTGTGGCTTCCTGCTGCTGCTGATGAAGAG 1175
Qy 361 ALAProGluLLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuLLeuLLeuLLeu 380
Db 1176 GCTCCTGAGATCTCAAGTGGCTCCCAAGATCCCTTGCTCCATCTGATTAACAAACATC 1235
Qy 381 LeuHisLLeuProGlnLLeuLLeuHisLLeuValAlaAsnGluLysPheHisAspLysHis 400
Db 1236 CTGCACATCCCGCTCAGATTTGCACTTGAGCTATGATTAATCTTCATGACAAAGCAC 1295
Qy 401 SerLeuThrGluLLeuArgAspSerLeuLeuAapLeuGlyAapValPhePheValVal 420
Db 1296 TCCCTGACTGAATCCGAGACAGTCTTCTGACTTGCTGAGATGTGTCTTTGTGGTC 1355
Qy 421 ProAlaLeuLLeuThrAlaArgLysHisAspAlaGlyValAlaProValLysThrThrGlu 440
Db 1356 CCTGACATGATCAGCTGATGATCAGAGATGCTGGTGGACCTGTCTACTCTATGAG 1415
Qy 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
Db 1416 TTTCGGCACCCGGCTCAGTGTCTTAAGACAGAAAGCCGGCTTTGTCTCAAGCCCAACAC 1475
Qy 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAapLLeuValMetPhe 480
Db 1476 GCTGATGAAGTCCGCTTGTGTTCGATGGTGCCTTCGAAGGGGAGCATGTGTATGTTTC 1535
Qy 481 GlnGlyAlaThrGlnGlnGlnLysLeuSerArgLysMetMetLysThrTrpAlaThr 500
Db 1536 GAAGGAGCCAGAGAGAGAGAAATTAAGTGAAGCCGAGAGATGATAAATATACGGGCTACC 1595
Qy 501 PheAlaArgThrGlnLysProAngLysAsnAspLeuSerLeuTrpProAlaLysLeu 520
Db 1596 TTTCGTCAGAACCGGAAATCTTAATGGAACAGACTGTCTCTGTGCGCAGCTTAATCTG 1655
Qy 521 ThrGlnGlnLysLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Db 1656 ACTGAGCAGTACCTCAGCTGAGCTTGAAATGAGCTCCGACAGACTCAAAAGAACCG 1715
Qy 541 ArgValAspPheThrThrSerThrLLeuProLeuLLeuSerAlaSerAspMetLeuHis 560
Db 1716 CGGGGTGATTTTGGACAGCAGCACTCCCTGATCTGTCTGCTCCGACATGCTCCAC 1775
Qy 561 SerProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePhePheCysAla 580
Db 1776 AGTCTCTTTCTTCTTAACCTTCTCTCTCTCTCCAGCCTTTCTTTTCTTTGTGCT 1835
Qy 581 Pro 581
Db 1836 CCT 1838

```

```

; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 40
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-40

Alignment Scores:
Pred. No.: 0 Length: 1728
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservative: 3
Best Local Similarity: 97.02% Mismatches: 6
Query Match: 94.67% Indels: 8
DB: 18 Gaps: 1

US-10-023-515-2 (1-581) x US-10-451-168-40 (1-1728)
Qy 12 TrpCysPhePheLeuLLeuGlnProLeuLeuGlnHisArgGlnTrpGlyLysThrGly 31
Db 40 TGGGCTATCTGGGTCTTTCAGACCCCTC-----ACCAAAGG 75
Qy 32 ProSerAlaGlnGlyProGlnArgAsnThrArgLeuGlyTrpLLeuGlnGlyLysGlnVal 51
Db 76 CTTTCTGTAAGGCGCACAGAGAAACAGAGCTGGATGATTCAGAGGCAAGCAAGTTC 135
Qy 52 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 71
Db 136 ACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTCTTCGAGAGTCCCTTGTGCTGCC 195
Qy 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProLTrpAspAsnLeuArg 91
Db 196 CCGCTGGATCCCTCGCATTAACGAACCGGACGCTGATGCGCTCGGATTAACCTTGGCA 255
Qy 92 GluAlaThrSerLysProAsnLeuLysLeuGlnAsnSerGlnTrpLLeuLeuAspGln 111
Db 256 GAAGCAGCTCTCAATCTTAATTTGTGCTTCAAGAACTAGAGTGTCTCTTAATCA 315
Qy 112 HisMetLeuLysValHisLysTrpProLysPheGlyValSerGluAspCysLeuLysLeu 131
Db 316 CATATGCTCAAGTGCATTACCCGAAATTCGAGAGTCAAGAAAGCTGCTTACTGAAAC 375
Qy 132 ILLeuTrpAlaProAlaHisAlaAspThrGlySerLysLeuProValLLeuValTrpPhePro 151
Db 376 ATCTATGCGCTGCGCCAGCGATTAACAGGCTCAAGCTCCCGCTTGTGTGTGCCA 435
Qy 152 GlyGlyAlaPheLysThrGlySerAlaSerLLeuPheAspGlySerAlaLeuAlaLysTr 171
Db 436 GAGGTGCTTCAAGACTGCGCTCAAGCTTCTTAATGAGTGTGCGCTGCTGCTGCTGCT 495
Qy 172 GluAspValLeuValValValGlnLysArgLeuGlyLLeuPheGlyPhePheThrThr 191
Db 496 GAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555

```

```

RESULT 6
US-10-451-168-40
; Sequence 40, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168

```



```

Db 454 CGAAGACCACTCTACCTATTTGCTCCAGAACTCAGAGGCTGCTTACAT 513
Qy 111 GlnHsmetLeuValHisTyrProlysRheglValSerGluAspCysLeuTyrLeu 130
Db 514 CAACGATGCTCAAGGTGATTAACCGAAATTCGAGGTGCAGAAAGACGCTCTACCTG 573
Qy 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLeuProValLeuValTyrPhe 150
Db 574 AACATCTATGCGCTCCGCCACCCATACAGGCTCCAGCTCCCTGTTGGTGTTC 633
Qy 151 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
Db 634 CAGAGAGGCTCTCAAGCTGCTCAGGCTCATCTTGTATGGGTCCGCTGCTGCTG 693
Qy 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
Db 694 TATGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 753
Qy 191 ThrTyrAspGlnHisAlaProGlyAsnTyrAlaPheLeuAspGlyValAlaAlaLeuSer 210
Db 754 ACATGGGATCAGCATGCTCTCCGGGAACTGGGCTTCAGAGACCAAGTGGCTCTGTC 813
Qy 211 TyrValGlnTyrAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
Db 814 TGGGTCCAGAAAGACATGAGTCTTCGGTGGGACCCAGCTGTGTGACATCTTGGC 873
Qy 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLeuGlyLeu 250
Db 874 GAGTCCGGGGGAGCCATAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 933
Qy 251 PheHisValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 270
Db 934 TTCACAAAGCCATATGAGAGGTGGGGGAGGCTCATCTCTTACCTGAGGCCCAATGAT 993
Qy 271 TyrGlnTyrSerGluAspLeuGlnValAlaAlaHisPheCysGlyAsnAsnAlaSerAsp 290
Db 994 TATGAGAAAGATGAGAGCACTGAGGAGGTGTGCAATTTCTGTGTGTGTGTGTGTGT 1053
Qy 291 SerGlnAlaLeuLeuArgCysLeuArgThrTyrProSerTyrGluLeuLeuThrLeuSer 310
Db 1054 TCTGAGGCTCTGCTGAGGTGCTGAGGACAAACCTCCAGAGAGCTGTGACCTCTGAC 1113
Qy 311 GlnTyrThrTyrSerPheThrArgValValaAspGlyAlaPhePheProAsnGluProLeu 330
Db 1114 CAGAAACCAAGTCTTCTCACTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCTTA 1173
Qy 331 AspLeuLeuSerGlnTyrAlaPheLeuValaIleProSerIleIleGlyValaAsnHis 350
Db 1174 GATCTATTGCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGACTCAATAACAC 1233
Qy 351 GlnCysGlyPheLeuLeuProMetTyrGlnAlaProGluIleLeuSerGlySerAsnTyr 370
Db 1234 GAGTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1293
Qy 371 SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrTyrLeuHisLeu 390
Db 1294 TCCCTTGGCTCTCATCTGATTAACAAACATCTGCAATCTCCGCTCAGATTTGTGACCTT 1353
Qy 391 ValAlaAsnGlnTyrPheHisAspTyrHisSerLeuThrGlnIleArgAspSerLeuLeu 410
Db 1354 GTGGCTATTGATATCTTCATGACAAAGCACTCCCTGACATCGAAATCGAGACAGTCTTCG 1413
Qy 411 AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArg 430
Db 1414 GACTTGTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1473
Qy 431 AspAlaGlyAlaProValTyrPheTyrGlnPheArgHisArgProGlnCysPheGlnAsp 450
Db 1474 GATGCTGGGACCGCTGCTACTTCTATAGATTTCGACACGGGCTCAGAGCTTTTAAGAC 1533
Qy 451 ThrTyrProAlaPheValTyrAlaAspHisAlaAspGluValaArgPheValPheGlyGly 470

```

```

Db 1534 ACGAAGCCGCTTTTGTCAAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGCT 1593
Qy 471 AlaPheLeuTyrGlyAspIleValMetPheGlnGlyValaThrGlnGluGluTyrLeuLeu 490
Db 1594 GCTTCTCGAAGGGGAGACATGTATATGTTCGAAAGAGCCAGAGAGAGAAAGTTACTG 1653
Qy 491 SerArgTyrMetLeuTyrTyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
Db 1654 AGCCGGAAGATGATGAATATCTAGGCTACCTTGTCTGCAACCGGAAATCTTAATGGAAAC 1713
Qy 511 AspLeuSerLeuTyrProAlaTyrAsnLeuThrGlnTyrGlnIleLeuAspLeuAsn 530
Db 1714 GACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1773
Qy 531 MetSerLeuGlyGlnArgLeuLeuGluProArgValaAspPheThrSerThrIlePro 550
Db 1774 ATGAGCTTCGAGCAGAGACTCAAAAGACCGGGGTGGATTGTGACACAGCATTCCTCC 1833
Qy 551 LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 570
Db 1834 CTGATCTGTGTGCTCCGACATGCTCCACAGTCTCTTGTCTTAATCTTCTCTCT 1893
Qy 571 LeuLeuGlnProPhePhePhePheCysAlaPro 581
Db 1894 CTCTCCAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1926

RESULT 8
US-10-104-047-249
; Sequence 249, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1: full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-249

Alignment Scores:
Pred. No.: 0 Length: 2092
Score: 2877.00 Matches: 551
Percent Similarity: 98.40% Conservative: 2
Best Local Similarity: 98.04% Mismatches: 2
Query Match: 93.44% Indels: 7
DB: 17 Gaps: 2

US-10-023-515-2 (1-581) x US-10-104-047-249 (1-2092)
Qy 24 HisArgGlnTyrGlyTyrThrGlyProSerAlaGluGlyProGlnIleArgAsnThrArgLeu 43
Db 126 CACTTAACTGG-----GGGCTTCTGTGTGAAGGCCACAGAGAAACCAAGGCTG 176
Qy 44 GlyTyrIleGlnGlyLeuGlnValThrValLeuGlySerProValProValAsnValPhe 63
Db 177 GGATGGATTCAGGGCAGACAGTCACTGTGCTGGAGCCCTGTGCTGTGAACGTGTTTC 236
Qy 64 LeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsnProGlnPro 83
Db 237 CTGGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGCATTTTACCAACCCGAGCCT 296
Qy 84 AlaSerProTyrPheAsnLeuArgGlnAlaThrSerTyrProAsnLeu-----C 100
Db 297 GCATCGCCCTGGATTAACCTTGCAGAAACCACTCTCACTTAATTTGTAAAGACAGGTG 356
Qy 100 sLeuGlnAsnSerGluTyrLeuLeuAspGlnHisMetLeuTyrValHisTyrProTyr 120

```


QY 12 TRCYSPHSPheleuileuengInProleuLeuGlyNHiaArgGlnTrpGlyLysThrGly 31
| | | | | : : : : : | | | | | : : : : :
DB 40 TGGGATATCGGGTCTTGCAGACCC-----ACAAAGG 75
QY 32 ProSerAlaGluGlyProGlnArgAntHraGluGlyTrpIleGlnGlyLysGlnVal 51
| | | | | : : : : : | | | | | : : : : :
DB 76 CTTTCCTGGAAGGGCCAGAGAAACACAGAGCTGGATGATTCAGAGGCAAGCAAGTC 135
QY 52 ThrValIleuGlySerProValProValaenValPheLeuGlyValProPheAlaIlePro 71
| | | | | : : : : : | | | | | : : : : :
DB 136 ACTGTCGGAAGAGCCGTGCTGTAACGTGTTCTCGAGATCCCTTGTGCTGCTCC 195
QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 91
| | | | | : : : : : | | | | | : : : : :
DB 196 CGCGTGGATCCCTGGATTTAGAACCCGAGCCGTGATCCCTCGGATAACTTGGCA 255
QY 92 GluAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGlnTrpLeuLeuAspGln 111
| | | | | : : : : : | | | | | : : : : :
DB 256 GAAGCACTCTACCTAATTTGTGCTCCAGAACTCAGAGTGCGCTCTTAGATCAA 315
QY 112 HisMetLeuValHisTyrProLysPheGlyValaSerGluAspCysLeuTyrLeuAsn 131
| | | | | : : : : : | | | | | : : : : :
DB 316 CATATGCTCAAGGTGATTAACCGAAATTCGAGGTCCAGAAAGACTGCTTAACCTGAC 375
QY 132 IleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValIleuValTrpPhePro 151
| | | | | : : : : : | | | | | : : : : :
DB 376 ATCTATGCCCCCGCCACGCCCATACAGGCTCCAGACTCCCGTCTGGTGTGTTCCCA 435
QY 152 GlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleTyr 171
| | | | | : : : : : | | | | | : : : : :
DB 495 GGAGGTGCTTCAAGACTGCTCAGGCTCATCTTGATGGGTCCGCTGCTGCTGCTAT 495
QY 172 GluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 191
| | | | | : : : : : | | | | | : : : : :
DB 496 GAGGACGTGCTGTGTGCTGTCCAGTACCGGCTAAGAAATTTGGTTTCTTCCACCA 555
QY 192 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaIleuSerTrp 211
| | | | | : : : : : | | | | | : : : : :
DB 556 TGGATACAGACTGCTCCGGGAACTGGGCTTCAAGAGCAAGTGCTGCTGCTGCTG 615
QY 212 ValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 231
| | | | | : : : : : | | | | | : : : : :
DB 616 GTCCAGAAAGAACATGAGTTCTTCGTTGGGAGACCCAGCTCTGTGACCATCTTGGCGAG 675
QY 232 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPhe 251
| | | | | : : : : : | | | | | : : : : :
DB 676 TCCGGGGGAGCATAGTGTCTTCTAGTCTTATACGTCTCCATGGCCAAAGCTTATTC 735
QY 252 HisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyr 271
| | | | | : : : : : | | | | | : : : : :
DB 736 CACAAAGCATATGAGAGTGGGTGGCATCATCTCTTACCTGAGGCCCAATGATAT 795
QY 272 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 291
| | | | | : : : : : | | | | | : : : : :
DB 796 GAGAAAGAGTGAAGACCTGAGGTGGTGCATTTCTGTGGTAAACAATCGTCAAGCTCT 855
QY 292 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGln 311
| | | | | : : : : : | | | | | : : : : :
DB 856 GAGGCGCTGCTGAGGTGCTGAGCAAAACCTCCAGAGACTGTCGACCTCAGCCGAG 915
QY 312 LysThrLysSerPheThrArgValValaAspGlyAlaPhePheProAsnGluProLeuAsp 331
| | | | | : : : : : | | | | | : : : : :
DB 916 AAAACAAGTCTTTCACTCGAGGTGATGAGTGTCTTCTTCTTAATAGCTCTAGAT 975
QY 332 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValaAsnAsnIleGlu 351
| | | | | : : : : : | | | | | : : : : :
DB 976 CTATTGTCTCAAAAGCATTTAAAGAAATTCCTTCATCATCGAGTCAATTAACCAAGAG 1035
QY 352 CysGlyPheLeuLeuProMetLysGlu-----AlaProGlnIleLeuSerGlySerAsn 369
| | | | | : : : : : | | | | | : : : : :
DB 1036 TGTGCTTCTGCTGCTATGTGAAGAAATTCGCTGTCCATATGCCACTCCCTCAAC 1095

QY 370 LysSerLeuAlaLeu-----HisLeuIleGlnAsnIleLeuHisIlePro 384
| | | | | : : : : : | | | | | : : : : :
DB 1096 CGTGATGACGTTTGGCTTCAACAGCTGGGCAATTTCCACAGAAACATGACACATCCCG 1155
QY 385 ProGlnTyrLeuHisLeuValAlaAsnGlnTyrPheHisAspLysHisSerLeuThrGlu 404
| | | | | : : : : : | | | | | : : : : :
DB 1156 CCTCAGTATTGCACTGTGGCTAATGAAATCTTCCATGACAAACACCTCCGTGACTGAA 1215
QY 405 IleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIle 424
| | | | | : : : : : | | | | | : : : : :
DB 1216 ATCCAGACAGTCTTCTGAGCTTGTGAGATGATGTTCTTGTGCTCCCTGACTGATC 1275
QY 425 ThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArg 444
| | | | | : : : : : | | | | | : : : : :
DB 1276 AAGCTCGATATCACAGAGATCTGTGACCTGTCTACTTCTTAATGAGTTTGGCACCGG 1335
QY 445 ProGlnCysPheGluAspThrLysProAlaPheValValaAspHisAlaAspGluVal 464
| | | | | : : : : : | | | | | : : : : :
DB 1336 CCTCAGTGTCTTGAAGACAGAAAGCAGCTTTTGTCAAAAGCCGACACAGCTGATAGATC 1395
QY 465 ArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThr 484
| | | | | : : : : : | | | | | : : : : :
DB 1396 CGTTTGTGTGGTGTGCTCTTCTGAAGGGGACATGTTATGTTGCAAGAGCCACG 1455
QY 485 GluGluGluLysLeuLeuSerArgLysMetCysTyrTrpAlaThrPheAlaArgThr 504
| | | | | : : : : : | | | | | : : : : :
DB 1456 GAGGAGAGAAAGTTACTAGCGCCGAGATGATGAATATCTGGGCTTACTTCTGAAAC 1515
QY 505 GlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeuThrGluGlnTyr 524
| | | | | : : : : : | | | | | : : : : :
DB 1516 GCGAAATCTTAATGGGAAAGACCTGTCTGTGGCCAGCTTAATCTGACTGAGCAGTAC 1575
QY 525 LeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArgValaAspPhe 544
| | | | | : : : : : | | | | | : : : : :
DB 1576 CTCAGCTGAGCTTAACTAGAGCTTCCGACAGACCTCAAAAGAACCGCGAGAGTGTG 1635
QY 545 TrpThrSerThrIlePro 550
| | | | | : : : : : | | | | | : : : : :
DB 1636 TGGGTGACGGGTATCCT 1653

RESULT 10
US-10-094-749-736
Sequence 736, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKRO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094, 749
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350, 435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1

[illegible]

Db	1005	GAGGCCCTGTCAGTGGCTCTGAGACAAACCTCCAGAGACTGTGACCTCCAGCCAG	106
Qy	312	LYSTHTLYSSErPheThrArgValValAArgGlyAlaPhePheProAsnGluProLeuAsp	331
Db	1065	AAACCAAGCTTTCTCACTCGAGAGGTGATGGGTCTTTCTTATATGAGCTCTAGT	1122
Qy	332	LeuLeuSerGlnIlysaIaPheIlysaIaIleProSerIleIleGlyValAsnAsnIleGlu	351
Db	1125	CTATTGCTCGAAGAACATTTAAAGCAATTCCTTCATTAATCGAGTCAATACCAAG	1184
Qy	352	CysGlyPheLeuLeuProMetIlysgIlyAlaProGluIleLeuSerGlySerAsnIlySer	371
Db	1185	TGTGGCTCTCTGGCTGCTGAAGAGAGGCTCTGAGATCTCATGGCTCCACAAGATCC	1244
Qy	372	LeuAlaLeuHileuIleGlnAsnIleLeuHileuHileuIleProProGlnIlyLeuHileuVal	391
Db	1245	CTTGCCCTCATCTGTATCAAAACATCTCGACATCCCGCTCAAGTATTTCCACTTGCTG	1304
Qy	392	AlaAsnGlnIlyTyPheHileAspIlyAsnIlySerLeuThrGluIleArgAspSerLeuLeuAsp	411
Db	1305	GCTAAAGAAATCTTCATGACAGACACTCCCTGATGAAATCCAGACAGCTTCTTGAC	1364
Qy	412	LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHileAspAsp	431
Db	1365	TGTCTTGGAAGATGTGTCTTGTGTCTCCCTGCATCGATCAAGCTCGATATACAGAG---	1421
Qy	432	AlaGlyAlaProValIlyTyPheTyrGluPheArgHileAspProGlnCysPheGluAspThr	451
Db	1421	-----	1421
Qy	452	IlyPProAlaPheValIlyAlaPheHileAspGluValArgPheValPheGlyGlyAla	471
Db	1421	-----	1421
Qy	472	PheLeuIlyGlyAspIleValMetPheGluGlyAlaThrGluGluGluIlyLeuLeuSer	491
Db	1422	-----GAGAGACCCAGAGAGAGAGAGATTCAGAGC	1454
Qy	492	ArgIlyMetMetIlySerTyrTyrAlaThrPheAlaArgThrGlyIlyAspProAsnGlyAsnAsp	511
Db	1455	CGAAGAAAGAAAGAAATCTGGGCTACCTTTGCTGAAACCGGGAATCTTAATGGGAAGAC	1514
Qy	512	LeuSerLeuTyrProAlaTyrAsnLeuThrGlnIlyTyrLeuGlnIlyLeuAspLeuAsnMet	531
Db	1515	CTGTCTCTGTGGCCAGCTTAATTAATCTGAGACGAGTCTCCAGCTGGACTTGAACATAG	1574
Qy	532	SerLeuGlyGlnArgLeuIlyGluProArgValAspPheThrPheSerThrIleProLeu	551
Db	1575	AGCCTCGAGACAGACATCAAGAAACCGCGGTGGATTATTTGGACCAAGACACATCCCCCTG	1634
Qy	552	IleLeuSerAlaSerAspMetLeuHileSerProLeuSerSerLeuThrPheLeuSerLeu	571
Db	1635	ATCTGTCTGTGGCTCGACATGCTCCAAAGTCTCTTCTTCACTTAATCTTCTCTCTC	1694
Qy	572	LeuGlnProPhePhePhePheCysAlaPro	581
Db	1695	CTCCAGCTTCTTTTCTTTTGTGGCTCT 1724	
RESULTE 11			
US-10-114-270-195			
Sequence 195, Application US/10114270			
Publication No. US20040030110A1			
GENERAL INFORMATION:			
APPLICANT: Guo, Xiaojia			
APPLICANT: Kekuda, Ramesh			
APPLICANT: Miller, Charles E.			
APPLICANT: Malynkar, Uriel M.			
APPLICANT: Spytek, Kimberly A.			
APPLICANT: Paturajan, Meera			
APPLICANT: Liu, Ziaohong			
APPLICANT: Gusev, Vladimir Y.			
APPLICANT: Li, Li			


```
QY 424 ILeThraLaAArgTYrHIEARGAAsPaLaGlaYAlaProVaLTYrPheTYrGluPheArGHIS 443
Db 1273 ATCAAGCTCGATATTCACAGAGATGCTGTGCACTGTCTATGAGTTCGGGCAC 1332
QY 444 ArgProGlnCYsrPheGlaAerThrLYsrProLaPheVaLYsAlaAerPHIsAlaAspGlu 463
Db 1333 CGGGCTCAGTCTTTGAAAGACACAGACGACCTTTTGCATCAAGCCGACCGCTGATGA 1392
QY 464 ValArgPheValAPheGlyGlyAlaPheLeuLYsrGlyAspGlyLeValMetPheGluGlyAla 483
Db 1393 GTCCGCTTTGTGTTCGGTGTGTGCTCTCTGAAAGGGGACATGTGTATGTTGGAAGGAGCC 1452
QY 484 ThrGluGluGluLYsrLeuLeuSerArgLYsrMetCysTYrTrpAlaThrPheAlaArg 503
Db 1453 ACCGAGAGAGAGAGAACTACTGAGCCGGAAGATGATGAATAACTGGGCTTACTTGTCTCA 1512
QY 504 ThrGlyAsnProArgnGlyAsnAspLeuSerLeuTYrProLaTYrAsnLeuThrGluGln 523
Db 1513 ACCGGAAATCTTAATGGGAAAGACCTGTCTGTGGCAGCTTATATCTGACTGAGCAG 1572
QY 524 TYrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLYsrGluProArgValAsp 543
Db 1573 TACCTCAGCTGGAATCTTGAACATGAGGCTCGACAGACTCAAGAAACCGGAGAGAT 1632
QY 544 PheTrpThrserThrIlePro 550
Db 1633 GTGTGGGTGACGGGGTATCTCT 1653

RESULT 12
US-10-381-898-20
; Sequence 20, Application US/10381898
; Publication No. US2004008687A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKY, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIPFIN, Jennifer A.;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
; APPLICANT: AKVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANMALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junning; YAO, Montique; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ. ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473645CB1
US-10-381-898-20

Alignment Scores:
Pred. No.: 1,516-300 Length: 1857
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 84.07% Indels: 23
DB: 18 Gaps: 1

US-10-023-515-2 (1-581) x US-10-381-898-20 (1-1857)

QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLeuGln 50
Db 274 GGGCTTCTGCTGAAAGGCGCACAGAGAAACAGGCTGGAGATGATTCAGGGCAACAA 333
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
Db 334 GTCACTGTGCTGGAGAACCTCTGCTGTGTGAACGTGTTCTCGGAGTCCCTTTGCTGCT 393
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProLaSerProTTPAspAsnLeu 90
Db 394 CCCCCTGGGATCCCTGCGATTTACGAAACCGAGCTGACATCGCTGGGATTAACCTTG 453
QY 91 ArgGluAlaThrSerTYrProAsnLeuCYsrLeuGlnAsnSerGluTrpLeuLeuAsp 110
Db 454 CGAAGAACCACTCTCCATCCCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGAT 513
QY 111 GluHisMetLeuLYsrValHisTYrProLYsrPheGlyValSerGluAspCYsrLeuTYrLeu 130
Db 514 CAATATATGCTCAAGGTGATTCACCCGAAATTCGGAGTCTCAGAAAGATGCTCTACTG 573
QY 131 AsnIleTYrAlaProLaHisAlaAerThrArgLYsrLeuLeuProValIleTrpPhe 150
Db 574 AACATGTATCGGCTGCCACCGCCGATACAGGCTCCAACTCCCTCCCTTGGTGTGTC 633
QY 151 ProGlyGlyAlaPheLYsrThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
Db 634 CAGAGAGGTGCTTCAAGACTGTGCTCAGCTCCATCTTGAATGGGTCCGCTGTGCTGCC 693
QY 171 TYrGluAspValLeuValValValGlnTYrArgLeuGlyIlePheGlyPhePheThr 190
Db 694 TATGAGAGAGTGTGTTGTGTGTGCTGCTGACGATCCGCTAGGAATATTTGGTTCTTCA 753
QY 191 ThrTPAspGlnHisAlaProGlyAsnTrpAlaPheLYsrAspGlnValAlaLeuSer 210
Db 754 ACATGGATCAGACATGCTCCGGGAAATGGGCTTCAAGAACAGATGGTGTCTGTCTC 813
QY 211 TrpValGlnLYsrAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
Db 814 TGGGTCCAGAAAGAACATGCAATGCTTCCGTGGGAGCCCACTGTGTACCATCTTTGGC 873
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLYsrLeu 250
Db 874 GAGTCGGCGGAGCATTAAGTGTTCATGCTTATACGTCTTCCATGCGCAAGGCTTA 933
QY 251 PheHisLYsrAlaIleMetGluSerGlyValAlaIleIleProTYrLeuGlnAlaHisAsp 270
Db 934 TTCCAAAAGCCATCATGAGAGAGTGGGTGCGCATCATCTTCACTGAGGCCCATGAT 993
QY 271 TYrGluLYsrSerGluAspLeuGlnValValAlaHisPheCYsrGlyAsnAsnAlaSerAsp 290
Db 994 TATGAGAGAGTGAAGACCTCAGTGTGTTGCAATTTCTGTGTAACATGCTTCAGAC 1053
QY 291 SerGluAlaLeuLeuArgCYsrLeuArgThrLYsrProSerLYsrGluLeuThrLeuSer 310
Db 1054 TCTGAGGCTTGTGAGGTGCTGAGGACCAAAACCTTCAAGAGGTGCTGACCTTACG 1113
QY 311 GlnLYsrThrLYsrSerPheThrArgValValAspGlyValaPhePheProAsnGluProLeu 330
```

Db	1114	CAGAAAACAAAGCTTTTCACTCGAGTGGTGAATGGTCTTTCTTCTTAATGAGCTCTA	11173
Qy	331	AspleuleuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHis	350
Db	1174	GATCATATATGTCACAAAGACATTTAAACAATTCCTTCATCATTCGGAGTCATTAACAC	12333
Qy	351	GluCyGgLYPheLeuLeuProMetLysGlnLaProGluIleLeuSerGlySerAsnLys	370
Db	1234	GAGGTGGCTTCTGCTGCTTAATG-----	1257
Qy	371	SerLeuAlaLeuHisIleLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisIleu	390
Db	1258	-----CACATCCGGCCCTCAATATTGGACCTT	1284
Qy	391	ValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeu	410
Db	1285	GTGGCTAATGAATATCTTCATGACACAGACATCCCTGACTGAATCCGAGACAGCTTCTG	13444
Qy	411	AspleuleuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArg	430
Db	1345	GACTTGCTTGGAGATGGTGTCTTGTGGTCCCTGACATGATCACAGTTCGATATCACGA	1404
Qy	431	AspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAsp	450
Db	1405	GATGCTGTGTCACCTGCTTACTTCTATAGATTTCCGACACCGGCTCACTGCTTGAACAC	1466
Qy	451	ThrLysProAlaPheValLysAlaAspHisIleAspGluValArgPheValPheGlyGly	470
Db	1465	ACGAAGCCAGCTTTGTGTCAAGCCGACACCGCTGATGAAGTCCGCTTGTGTGGTGGT	1524
Qy	471	AlaPheLeuLysGlyAspIleValMetPheGluGlyValaThrGluGluGluLysLeuLeu	490
Db	1525	GCCCTTCCGAAAGGGGAGCATTTGTTATGTTCAAGGAGCCAGAGAGAGAAAGATTACTG	15844
Qy	491	SerArgLysMetCysMetLysTyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn	510
Db	1585	AGCCGGAAGATGATGAATACTGGGCTACCTTTGCTCAACCGGGAATCTTAATGGAAC	16444
Qy	511	AspleuSerLeuTyrProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspleuAsn	530
Db	1645	GACCTGTCTGTGGCCAGCTTATATATCTGACTGAGCAGTACCTCCAGCTGACTGAAC	1704
Qy	531	MetSerLeuGlyGlnArgLeuLysGluProArgValAspPheTyrPThrSerThrIlePro	550
Db	1705	ATGAGCCCTCGGACAGACATCAAAAGAACCGGAGAGATGTGTGGTGAACGGAGTATCTT	17644
RESULT 13			
US-10-233-933A-3			
Sequence 3, Application US/10233933A			
Publication No. US2004021417A1			
GENERAL INFORMATION:			
APPLICANT: Miyazaki, Masaao			
APPLICANT: Yamashita, Tetsuo			
TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER			
FILE REFERENCE: SHIG FP02US06			
CURRENT APPLICATION NUMBER: US/10/233, 933A			
CURRENT FILING DATE: 2002-09-03			
PRIOR APPLICATION NUMBER: JP2002-057908			
PRIOR FILING DATE: 2002-04-03			
NUMBER OF SEQ ID NOS: 4			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 3			
LENGTH: 2145			
TYPE: DNA			
ORGANISM: Felis catus			
FEATURE:			
NAME/KEY: CDS (175)..(1803)			
FEATURE:			
NAME/KEY: misc feature			
LOCATION: (1804)..(2145)			
OTHER INFORMATION: n is a, c, g, or t			
US-10-233-933A-3			

Alignment Scores:			
Pred. No.:	2.5e-231	Length:	2145
Score:	2015.00	Matches:	388
Percent Similarity:	80.87%	Conservative:	60
Best Local Similarity:	70.04%	Mismatch:	102
Query Match:	65.44%	Indels:	4
DB:	20	Gaps:	2
US-10-023-515-2 (1-581) x US-10-233-933A-3 (1-2145)			
QY	31	GIYProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPILeGlnGlyArgGln	50
DB	238	GGGCGACGCTGCTGATGCACAGTAAAGGAGCACAGGCTGGATGGGTCGGGGAGAAAGCAA	297
QY	51	ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIa	70
DB	298	ACCACTGACTACGGAAAGACCGTGCCTGTGAACATGTCCTCGGGATCCCTATCTGCA	357
QY	71	ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu	90
DB	358	CTCTCTCTAGGGCCCCCTCCGATTTTAAAGAACCAAGCTGCTTCGCCGGGAATGACTTC	417
QY	91	ArgGlnAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGluTyrPheLeuLeuAsp	110
DB	418	CGAATGCGCATCTCACTCCATAATTATGCTTCAGAGACTTAAGATGGCTGGCTCTCAT	477
QY	111	GlnHisMetLeuValHisTyrProArgPheGlyValSerGlnAspCysLeuTyrLeu	130
DB	478	CAACACGTTCTCAAGATGCGTTACCCCAATTTGAAAGCGTCCGAAGACTGCTGTAACCT	537
QY	131	AsnIleTyrAlaProAlaHisAlaAspThrGlySerIleuProValLeuValTyrPhe	150
DB	538	AAACTCTATGGCCAGCCCAAGCCGAGCAATGGCTCCAACTCCCTGTATGATGTGTTC	597
QY	151	ProGlyGlyAlaPheIleThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa	170
DB	598	CCCCGGGGGTGCTTCAAGATGGGCTGACCTTCCTTCGATGGGTCGGCTTGCTGCC	657
QY	171	TyrGlnAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr	190
DB	658	TACAGAGACGGCTGATGCTGACTACCAACCAAGACCGCTTGAAGAAATTTGGTTTTCAC	717
QY	191	ThrTyrAspGlnHisAlaProGlyAsnTyrAlaPheIleuAspGlnValAlaLeuSer	210
DB	718	ACAAGGGATGAGCATGCGCCGGGGAGTGGGCTTCGTCGAGCAAGTGGCTGCTTAAC	777
QY	211	TyrValGlnIleuAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly	230
DB	778	TGGGTCCGGCAACATGCAATGCTTCTCGTGTGTGACCAACCTCGTGACCATCTTTGGA	837
QY	231	GlnSerIleGlyAlaIleSerValSerSerIleuIleuSerProMetAlaIleGlyLeu	250
DB	838	GAGTCAGGGAGGAGCATCAGTTCACGCTCATCTGTCCCAATGACCAATGGCTTA	897
QY	251	PheHisIleValIleMetGlnSerGlyValAlaIleIleProTyrLeu---GlnAlaHis	269
DB	898	TTCCACAAAGCCATCATGAGAGATGGGGTGGCCATCCGCTTTACTATAGAACCCCT	957
QY	270	AspTyrGlnIleuSerGlnAspPheGlnValAlaHisIlePheCysGlyAsnAsnAlaSer	289
DB	958	GGTATGAGAGGAAGAAAGATTTTCAGTGTCTCGCTATCTGTGTGTCATGCTGCT	1017
QY	290	AspSerGlnAlaLeuLeuAspCysLeuArgThrTyrProSerIleuGlnLeuLeuThrLeu	309
DB	1018	GACTCTGCTGCCCTGCTGCACTGCTGAGGGCAAAACCTCCAGAGACTTATGACATTC	1077
QY	310	SerGlnIleuThrIleSerPheThrArgValValAlaAspGlyAlaPhePheProAsnGluPro	329
DB	1078	AGCAAGAAACTCACGTTTTCATCCAGTATGATGACTTTTCTTTCTTCATGAGACT	1133
QY	330	LeuAspLeuLeuSerGlnIleuValPheIleValIleProSerIleIleGlyValAlaAsnAsn	349

```
Db      1138 GTAGCCCTATTGACCAAAAAGCATTTAATTCACTTCCTTATCATCGAGTCATAATAC 1197
Qy      350 Htsglucyvglypneuleuprometylsglualaprogluileuserglyserasn 369
Db      1198 CACGAGTGTGCTTCCTCTG--TCCAGGAGTTTTCTGAGATCTCGGCGCTCCAC 1254
Qy      370 lyaSerleuAlaLeuHsleuileglnlanilLeuHsileuHsileuProProglintyLeuHs 389
Db      1255 AGGCTCTGGCCCTCTAGTATGACACAGCTCTGAAATTTCCACCCAGTATTTGAC 1314
Qy      390 LeuValAlaangluTyPheHsAaPlyHsileuSerleuthrgluilearAaPserleu 409
Db      1315 CTGTGGCTAGTATTACTTCTACACACACACTCCCGGTTGAATAACAGATAGTTT 1374
Qy      410 LeuAaPleuLeuGlyAaPValPhePheValProAlaLeuileHraAlaGlyTyrHs 429
Db      1375 CTGACCTGCTTGGAAATGTCTCTTTGTGTCTGGGTGTCAAGTCATATCAT 1434
Qy      430 AaGAPAlaGlyAaProValTyPheTyTgUpheAaHsAaPProglncysPheglu 449
Db      1435 AGAGATGTGTGACCTGTCTACTTCTATGAGTTTCAACACCCGCCCAAGTCTTAAC 1494
Qy      450 AaPThlyAaPProAlaPheValIysAlaAaPNIAlaAaPGLValAaPheValPheglu 469
Db      1495 GACACAGGCCCACTTCTGMAAGCCGATCACTGATGAATCCGCTTCTTTGGA 1554
Qy      470 GlyAlaPheleuLeuGlyAaPNIleValIlePhegluGlyAlaThrgluileu 489
Db      1555 GGTGCTTCTCTGMAAGCGCATTTGTCATGTTGMAAGAGCCACCGAGAGAGAAATTG 1614
Qy      490 LeuSerAaPlyMeMetleuTyTTPAlaThrPheAlaAaGThrglyAaPProangly 509
Db      1615 CTGAGAGGAAAGATGATGAGTACTGGGCCAACTTGTCTCGACCGGGGAACTTAACGGG 1674
Qy      510 AaAaPleuSerleuTyPProAlaTyAaPleuThrgluileuThrgluileuSerleu 529
Db      1675 GAAGGTGCTCTGTGGCCACCTACACCCAGACGAGCACTAGAGCTGAATTG 1734
Qy      530 AaMetSerleuGlyGlnAaPleuLeuGluProAaPValAaPheTyPThSerThrlle 549
Db      1735 AGTGAAGGTGGACAGAACTGAAGGCAAGAGAGTGGAGTTTGGATGAATCCATT 1794
Qy      550 --ProleuileuSerAlaSerAaPMeleuHsleuSerProleuSerleuthrPhe 569
Db      1795 GTCCCTGATACCCCCCACTTCAGGCGCTCCCAAGTCTCTTNTCCCTTACTCTCC 1854
Qy      569 euSerleuLeuGlnProPhePhePhePheCyAaLaPro 581
Db      1855 TTCTTTGCTCCCGCTGCTTNTTTTTCGTGCTCA 1892
```

RESULT 14

```
US-10-233-933A-1
; Sequence 1, Application US/10233933A
; Publication No. US20040214171A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Masao
; APPLICANT: Miyazaki, Masao
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG PRO2US006
; CURRENT APPLICATION NUMBER: US/10/233, 933A
; PRIOR FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: JP2002-057908
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Felis catus
US-10-233-933A-1
```

Alignment Scores:

```
Pred. No.: 6.Se-228 Length: 1629
```

```
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Percent Local Similarity: 71.73% Mismatches: 85
Query Match: 64.47% Indels: 2
DB: 20 Gaps: 2
```

US-10-023-515-2 (1-581) x US-10-233-933A-1 (1-1629)

```
Qy      31 GlyProSerAlaGluGlyProGlnAaPNIleAaPNIleAaPNIleGlnGlyGln 50
Db      64 GGGCCAGCTGGTGAATGACACAGTGAAGAGACACAGCTGGATGGCTCCGGGAGAGAA 123
Qy      51 ValThrValleuGlySerProValProValAaPNIlePheleuGlyValProPheAla 70
Db      124 ACCACTGTAAGTGGAGAGACCGCTGTGTGAACATGTCTCGGAGATCCCTATGTCGA 183
Qy      71 ProPleuGlySerleuAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIle 90
Db      184 CCTCTCTAGGAGCCCTGCGATTTAAGCAACAAAGCTGCTGCGCGGAGATGACTTC 243
Qy      91 ArgGluAlaThrSerTyPProAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAa 110
Db      244 CGAATGCGCATGCTACCTTAATTAATGCTTCAGAGACTTAAGATGGCTGTCTCTAT 303
Qy      111 GlnHsileuLeuValHsleuTyPProAaPNIleAaPNIleAaPNIleAaPNIleAaPNIle 130
Db      304 CAACAGCTTCAAAAGTCCGTTACCCCAATGGAGACCGTCCGAGACCTGCTGACTT 363
Qy      131 AaNIleTyAaPProAlaNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAa 150
Db      364 AACATCTATGGCGCCAGCCCAATGGAGCAATGGCTCCCAACTCTCTGTCATGTGTGTC 423
Qy      151 ProGlyGlyAlaPheTyPThrglySerAlaSerleuAaPNIleAaPNIleAaPNIleAa 170
Db      424 CCGGAGGTGCTCAAGATGGCTGAGCTCTCTCTCATGAGTGGCTGGCTGGCTGGCT 483
Qy      171 TyTgUpAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAa 190
Db      484 TAGAGGACCTGCTGATCGTGAATCCAGTACCCAGTACCCAGTACCCAGTACCCAGTAC 543
Qy      191 ThrTyPProGlnHsileuAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIle 210
Db      544 ACAAGGATGAGCATGCCCGGAGAACTGGCTGCTGACCAAGTGGCTGCCCTTAAC 603
Qy      211 TyTgUpAlaNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAa 230
Db      604 TGGGTCCGGAACAATCGATGTTCTGAGTGAACCAAGCTCCGTCATCTTTTGA 663
Qy      231 GluSerAlaGlyAlaIleSerAlaSerleuileuSerProleuAaPNIleAaPNIleAa 250
Db      664 GAGTCAGCGGAGCCATCAAGTGTTCAGCTCATTTCTCCCAATGACCAATGCTTA 723
Qy      251 PheHsileuAlaIleMetleuSerGlyValAlaIleIleProTyLeu--GluAlaHs 269
Db      724 TTCACAAAGCCATCATGAGAGTGGGAGCACTGCTTATCATGATGAGACCCCT 783
Qy      270 AaPtyTgUpSerleuAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAa 289
Db      784 GGTATGAGAGAGAAAGATTTGCAAGTGTCTGGCGGTATGTGTGAGTGGATGCT 843
Qy      290 AaPserGluAlaLeuLeuAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIle 309
Db      844 GACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
Qy      310 SerGlnTyPThlySerPheThraGlyAlaAaPNIleAaPNIleAaPNIleAaPNIleAa 329
Db      904 AGCAAGAACTCACTTTTCCATTTCCAGTGAATGATCTTTTCTTTCGATGAGCCT 963
Qy      330 LeuAaPNIleuSerGlnTyAaPNIleAaPNIleAaPNIleAaPNIleAaPNIleAaPNIle 349
Db      964 GTAGCCCTATGACTCAAAAAGCATTTAATTCAGTCTCTTATCAACGAGTCAATTAAC 1023
Qy      350 Htsglucyvglypneuleuprometylsglualaprogluileuserglyserasn 369
```

```
|||||
Db 1024 CACGAGTGTGCTTCTTCTG---TCCACGAGGTTTCTGAGATCTCGGGGGCTCCAC 1080
Qy 370 LySerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProlGlnTyrLeuHis 389
Db 1081 AGGCTCTGGGCTTCTTACTTACTAGTACACGTTCCGAAATATTTCCACCCAGATATTTGAC 1140
Qy 390 LeuValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeu 409
Db 1141 CTGTGGCTGATCTACTTACTTCTACACAAAGCACTCCCGGTTGAATTCGAGATGTTT 1200
Qy 410 LeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHis 429
Db 1201 CTGAGCTTGTGAGATGATGTGCTTTGTGGTCCCTGGGGGTGTGACAGCTGATATCAT 1260
Qy 430 ArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGlu 449
Db 1261 AGAGATGCTGGTGCACTGTCTACTTCTATGATGATTCACACCCGCCAGTGTAAAC 1320
Qy 450 AspThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGly 469
Db 1321 GACACGAGGCGCACTTGTGTAAGCCGATCATCTGATGAATCCGCTTGTCTTTGGA 1380
Qy 470 GlyAlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeu 489
Db 1381 GGTGCTTCTCGAAGGCGCATTTGTGATGTTGGAAGGCGCACCGAGAGGAATTTG 1440
Qy 490 LeuSerArgLysMetMetLysTyrTrrAlaThrPheAlaArgThrGlyAsnProAsnGly 509
Db 1441 CTGAGCAGGAAGATGATGATGAGTACTGGGCGCAACTTGTCTGGAACCGGGGACCTTAAACGG 1500
Qy 510 AsnAspLeuSerLeuTrrProAlaTyrAsnLeuThrGluIleTyrLeuGluLeuAspLeu 529
Db 1501 GAAGGTGTGCTCTGTGGCGAGCTTACACCCAGACGAGACGATCTGAAGCTGATTTG 1560
Qy 530 AsnMetSerLeuGlyGlnArgLeuLysGlyLupProArgValAspPheTrrPheSerThrIle 549
Db 1561 AGTGTAGCGTGGGACAAATGTAAGAGCAAGAGGAGTTTGAATGAATTCATT 1620

RESULT 15
US-10-451-168-42
; Sequence 42, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1071
```

```
TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-42

Alignment Scores:
Pred. No.: 1,63e-203 Length: 1071
Score: 1780.00 Matches: 340
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.81% Indels: 0
DB: 18 Gaps: 0

US-10-023-515-2 (1-581) x US-10-451-168-42 (1-1071)
Qy 242 ILeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGluSerGlyValAla 261
Db 49 ATACTGTCTCCATGCGCCAAAGGCTTATTCACAAAGCATGATGGAAGGGGTGACC 108
Qy 262 ILeuProTyrLeuGluAlaHisAspTyrGluLysSerGluAspLeuGluValAla 281
Db 109 ATCATCCCTTACCTGAGGCCCATGATTAAGAAAGCTGAGACCTCGAGTGTGCA 168
Qy 282 HisPheCysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuArgCysLeuArgThrLys 301
Db 169 CATTTCTGTGGTAACAATGCGCTGAGACTGTAGGCGCTGAGGCTGAGGACAA 228
Qy 302 ProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValAlaAsp 321
Db 229 CCTCCAAAGAGCTGAGCCCTGACCCAGCAAAACAAAGCTTTTCATCTCGAGTGTGAT 288
Qy 322 GlyAlaPhePheProAsnGluProLeuAspLeuSerGlnLysAlaPheLysAlaIle 341
Db 289 GGTGCTTCTTCTTAAGAGCCTTAGATCTATGTGTGCAAAAGCATTTAAAGCATTT 348
Qy 342 ProSerIleIleGlyValAsnAsnHisGluCysGlyPheLeuLeuProMetLysGluAla 361
Db 349 CCTTCATCATCTGAGATGAATTAACCAAGATGTGGCTTCTGCTGCTATGAAAGAGGCT 408
Qy 362 ProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisIleGlnAsnIleLeu 381
Db 409 CCTGAGATCTTCAGTGGCTCAACAAAGCTTGTGCTTCATCTGATCAAAACATCTGTG 468
Qy 382 HisIleProProlGlnTyrLeuHisLeuValAlaAsnGluTyrPheHisAspLysHisSer 401
Db 469 CACATCCCGCTCAGATATTTGACCTTGTGCTATGATGATCTTCCAGACAGACATCTC 528
Qy 402 LeuThrGluIleArgAspSerLeuLeuAspLeuGluArgValPhePheValAlaPro 421
Db 529 CTGACTGAATCCGAGACAGTCTTCTGGACTTGTGAGATGTGTCTTGTGTGCTCCCT 588
Qy 422 AlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPhe 441
Db 589 GCACGTAGCACAGCTCGATATCACAGAGATCTGTGACCTGTCTACTTCTATGAGTTT 648
Qy 442 ArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHisAla 461
Db 649 CGGACCGGCTCAGTGTGTTGAAAGACAGAAAGCAAGCTTTGTCAAGCCGACCACTCT 708
Qy 462 AspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlu 481
Db 709 GATGAAGTCCGCTTGTGTGTGCGTGTGCTTCTGAAAGGGGACATGTTATGTTCGAA 768
Qy 482 GlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyrTrrAlaThrPhe 501
Db 769 GGAGCGACGAGGAGGAAGATTACTGAGCGGGAAGATGAATATCTGAGCTACCTTT 828
Qy 502 AlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrrProAlaTyrAsnLeuThr 521
Db 829 GCTGGAACCGGGAATCTTAATGGAGCAAGCTGTCTGTGTGCGCAAGCTTAATCTGACT 888
Qy 522 GluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArg 541
Db 889 GAGCAGTACTTCAGCTGAGCTTGAACATGAGCTTCGAGACAGAGCTCAAGAAACCGCGG 948
```

Qy	542	VaIAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAapMetLeuHisSer	561
Db	949	GTGGATTTTGGACGACCAATCCCCCTGATCTGTGCTCCGACATGCTCCACAGT	1008
Qy	562	ProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAlaPro	581
Db	1009	CCCTCTTCTTCCCTTAACTTCTCTCTCTCCCTCCAGCCTTCTTTCTTTGTGCTCCT	1068

Search completed: June 15, 2005, 15:44:59
Job time : 994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 08:55:45 ; Search time 87 Seconds
(without alignments)
2582.848 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079

Sequence: 1 MPGLTSSASQWCFLLIQP.....PLSSLTFLSLQPPFFFCAP 581

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq.16Dec04:*

1: Genesegp1980s:*\n2: Genesegp1990s:*\n3: Genesegp2000s:*\n4: Genesegp2001s:*\n5: Genesegp2002s:*\n6: Genesegp2003as:*\n7: Genesegp2003bs:*\n8: Genesegp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	581	5	ABR79537 Human car
2	3079	100.0	581	8	ADQ89094 Human uro
3	2915	94.7	575	5	ABP61004 Novel hum
4	2900	94.2	642	5	AAE25025 Human dtr
5	2900	94.2	642	8	ABM84114 Human dia
6	2607.5	84.7	581	5	ABP61005 Novel hum
7	2606	84.6	525	6	ADA54807 Human pro
8	2597.5	84.4	581	6	ABU54639 Human NOV
9	2586.5	84.1	618	5	ADRI9663 Human dtr
10	2454	79.7	469	7	ADB64065 Human pro
11	1985	64.5	542	8	ADP50145 Cat cauxi
12	1985	64.5	542	8	ADP50147 Cat cauxi
13	1780	57.8	356	5	ABP61006 Novel hum
14	1262.5	41.0	561	5	AAE20911 Rat carbo
15	1230	39.9	549	3	AAE58981 Breast an
16	1230	39.9	550	8	ADG98219 Human int
17	1230	39.9	559	4	AAAB31700 Protein e
18	1230	39.9	559	6	ABU53223 Human met
19	1230	39.9	559	6	ABR82851 CES2 rela
20	1230	39.9	559	7	AAE33960 Human car
21	1230	39.9	559	7	ADD46521 Human pro
22	1221	39.7	306	5	AAE20909 Human car
23	1213.5	39.4	561	7	ADD46519 Rat Prote
24	1191	38.7	607	4	ABU53222 Human met
25	1186.5	38.5	554	8	AD179890 Mouse liv

26	1181	38.4	583	4	ABG10273 Novel hum
27	1179.5	38.3	571	3	AAV71107 Human Hyd
28	1179.5	38.3	571	4	AAU12442 Human PRO
29	1179.5	38.3	571	4	AAE04101 Human gen
30	1179.5	38.3	571	5	AAU83696 Human PRO
31	1179.5	38.3	571	5	ABR84949 Human PRO
32	1179.5	38.3	571	5	ABG64341 Human alb
33	1179.5	38.3	571	6	ABO17886 Novel hum
34	1179.5	38.3	571	6	ABU69108 Human PRO
35	1179.5	38.3	571	6	ABU80843 Human PRO
36	1179.5	38.3	571	6	ABO33809 Novel hum
37	1179.5	38.3	571	6	ABU81140 Human sec
38	1179.5	38.3	571	6	ABO19424 Human sec
39	1179.5	38.3	571	6	ABU66840 Human PRO
40	1179.5	38.3	571	6	ABU59921 Novel sec
41	1179.5	38.3	571	6	ABU69085 Human PRO
42	1179.5	38.3	571	6	ABO25111 Human sec
43	1179.5	38.3	571	6	ABU82152 Novel hum
44	1179.5	38.3	571	6	ABU67116 Human sec
45	1179.5	38.3	571	6	ABU81549 Human sec

ALIGNMENTS

RESULT 1

ID ABR79537 standard; protein; 581 AA.

AC ABR79537;

DT 23-SEP-2002 (first entry)

DB Human carboxylesterase family member 53010.

KW Carboxylesterase; enzyme; human; analgesic; nootropic; antiinflammatory; diagnosis; therapy.

KM Homo sapiens.

XX Homo sapiens.

FH Key Location/Qualifiers

FT Peptide 1..26 /label= Signal_peptide

FT Protein 27..581 /label= Mature_protein

FT Domain 44..545 /note= "Carboxylesterase domain"

FT Region 125..135 /note= "predicted carboxylesterase type-B serine active site"

FT Active-site 219..234 /note= "predicted carboxylesterase type-B signature"

WO200250256-A2.

27-JUN-2002.

18-DEC-2001; 2001WO-US049075.

18-DEC-2000; 2000US-0256369P.

28-MAR-2001; 2001US-0279508P.

(MILL-) MILLENNIUM PHARM INC.

Curtis RAJ, Silos-Santiago I;

WPI; 2002-547936/58.

N-PSDB; ABN84302.

53010 nucleic acids, useful for diagnosing and treating e.g. vascular diseases, autoimmune diseases, or neurodegenerative diseases; as surrogate markers, in tissue typing and chromosome mapping.

Db 121 FGVEDCLYLIYAPAHADTGSKLPLVYVFGGAFTGSASIFDGSALAAYEDVLVVVVQ 180
 QY YRLGIFGPEPTTDDOAPGNWMAFKDOVAALSWQXNIIEFGGPPSSVTTFGESAGAISSVS 240
 Db 191 YRLGIFGPEPTTDDOAPGNWMAFKDOVAALSWQXNIIEFGGPPSSVTTFGESAGAISSVS 240
 QY 241 LILSPMAKGLPFKAIMESGVAIIPLYEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRT 300
 Db 241 LILSPMAKGLPFKAIMESGVAIIPLYEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRT 300
 QY 301 KPSKELLTISQKTKSFTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKE 360
 Db 301 KPSKELLTISQKTKSFTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKE 360
 QY 361 APEILSGNSKSLALHILQNIILHIIPPOYLILVANEXYFHDGHSLEIRDSLLDLIDGVFPV 420
 Db 361 APEILSGNSKSLALHILQNIILHIIPPOYLILVANEXYFHDGHSLEIRDSLLDLIDGVFPV 420
 QY 421 PALITARYHRDAGAVYFYEFPNRPQCEFDTPKPAFVKADHADEVFVEGAFKADIVWF 480
 Db 421 PALITARYHRDAGAVYFYEFPNRPQCEFDTPKPAFVKADHADEVFVEGAFKADIVWF 480
 QY 481 EGATEEEKLSHRKMKYVATFARTGNPNGNDSLMPAYNLTEOYLQDLNMSLGRLKEP 540
 Db 481 EGATEEEKLSHRKMKYVATFARTGNPNGNDSLMPAYNLTEOYLQDLNMSLGRLKEP 540
 QY 541 RVDFTSTITPLILASDMLHSPSLTFLSLQPPFFCAP 581
 Db 541 RVDFTSTITPLILASDMLHSPSLTFLSLQPPFFCAP 581

RESULT 3

ABP61004
ABP61004 standard; protein; 575 AA.

AC ACP61004;

DT 10-SEP-2002 (first entry)

DE Novel human protein. SEQ ID 91.

Human; cytosolic; vulnary; antiarteriosclerotic; antiparkinsonian;
 neurotropic; neuroprotective; immunosuppressive; haemostatic;
 antiinflammatory; cardiant; antilicer; vaccine; cancer; infection;
 cerebroprotective; anorectic; metabolic; Parkinson's disease;
 wound healing disorders; atherosclerosis; Parkinson's disease;
 Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 inflammation; neoplastic disease; nervous system disorder;
 cardiovascular disorders; pancreatitis; respiratory disorder;
 hyperproliferation; systemic autoimmune disease; hyper-immunity;
 developmental abnormality; gastrointestinal ulceration; neuropathy;
 haematological disease; metabolic disease; sperm dysfunction;
 thyroid disorder; hypothyroidism; brain damage; colitis;
 cone photo-transduction deficiency; neurological disease; stroke;
 anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 trachea; thymus; lymph node; muscular system; obesity; anorexia;
 growth abnormality; precocious puberty.

OS Homo sapiens.

PN MO200250105-A1.

PD 27-JUN-2002.

PF 17-DEC-2001; 2001WO-US049232.

PR 19-DEC-2000; 2000US-0256710P.

PR 20-DEC-2000; 2000US-0257048P.

PR 09-JAN-2001; 2001US-0260482P.

PR 30-JAN-2001; 2001US-0264922P.

PR 06-FEB-2001; 2001US-0266797P.

PR 19-MAR-2001; 2001US-0276988P.

PR 04-APR-2001; 2001US-0281535P.

PR 08-MAY-2001; 2001US-0289622P.
 XX (SMK) SMITHLINE BECHAM CORP.
 PA (SMK) SMITHLINE BECHAM PLC.
 PA (GLAX) GLAXO GROUP LTD.
 XX Agawall P, Birkeland M, Cogswell JP, Kabnick KP, Lai Y;
 PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
 DR MFI, 2002-508784/54.
 XX N-PSDB; AB086169.
 PT Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX Claim 1(a); Page 312-313; 335pp; English.

The invention relates to an isolated polypeptide with signal sequences
 which allow it to be secreted extracellularly or membrane associated. The
 activity of polypeptides of the invention may be described as,
 CC cytosolic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiant, antilicer, vaccine, antithyroid, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder,
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, anglogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention
 CC XX

Sequence 575 AA;

Query Match 94.7%; Score 2915; DB 5; Length 575;
 Best Local Similarity 97.0%; Pred. No. 2.3e-263;
 Matches 553; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 12 WCFPLILQPLIGHRQWGTGSAEGPQNTRLGWIQKQVTVLGSPPVNVFLGVPFAP 71
 Db 14 MAIWTLAP-----TKGSABGPQNTRLGWIQKQVTVLGSPPVNVFLGVPFAP 65
 QY 72 PLGSRFTNPQAPASFWNDLREATSYPNICLQNSEWLLIDQMLKHYRFGVSEDCLYLN 131
 Db 66 PLGSRFTNPQAPASFWNDLREATSYPNICLQNSEWLLIDQMLKHYRFGVSEDCLYLN 125
 QY 132 IYAPAHADTGSKLPLVYVFGGAFTGSASIFDGSALAAYEDVLVVVVQYLGIFGFTT 191
 Db 126 IYAPAHADTGSKLPLVYVFGGAFTGSASIFDGSALAAYEDVLVVVVQYLGIFGFTT 185
 QY 192 WDOHAPGNMAFKDOVAALSWQXNIIEFGGPPSSVTTFGESAGAISSVSLILSPMAKGLF 251
 Db 186 WDOHAPGNMAFKDOVAALSWQXNIIEFGGPPSSVTTFGESAGAISSVSLILSPMAKGLF 245
 QY 252 HKAIMESGVAIIPLYEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRTKPSKELLTISQ 311
 Db 246 HKAIMESGVAIIPLYEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRTKPSKELLTISQ 305
 QY 312 KTKSFTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKEAPBIIISGNSKS 371
 Db 306 KTKSFTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKEAPBIIISGNSKS 365
 QY 372 LALHLIQNIILHIIPPOYLILVANEXYFHDGHSLEIRDSLLDLIDGVFPVVALITARYARD 431

Db	366	LALHLIQNLINLIPQYLHLVANEYFHKHSLTEIRNDSLDDLGDFVEFVPALITARYRD	425
Qy	432	AGAPVYFPEFRHRPOCFEDTYPKAFVKADHAEVRFVPGAFPLKGDIVWFEGATEEKLIS	491
Db	426	AGAPVYFPEFRHRPOCFEDTYPKAFVKADHAEVRFVPGAFPLKGDIVWFEGATEEKLIS	485
Qy	492	RKMCKYMTFPAATGPNPNNDLSIMPAYNLTQOYLQDLINSLGRLKXPRVDEFTSTIPL	551
Db	466	RKMCKYMTFPAATGPNPNNDLSIMPAYNLTQOYLQDLINSLGRLKXPRVDEFTSTIPL	545
Qy	552	ILSASDMLHSPSSLTFLSLQLQPEFFFCAP	581
Db	546	ILSASDMLHSPSSLTFLSLQLQPEFFFCAP	575
RESULT 4			
ID	AAE25025	standard; protein; 642 AA.	
XX	AAE25025;		
XX	30-OCT-2002	(first entry)	
DB	Human drug metabolising enzyme (DME-10).		
XX			
KM	Human; drug metabolising enzyme; autoimmune; inflammatory disorder;		
KM	acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;		
KM	proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;		
KM	asthma; neurological disorder; Alzheimer's disease; Huntington's disease;		
KM	dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;		
KM	drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;		
KM	renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;		
KM	anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;		
KM	gastrointestinal disorder; gene therapy; virucide; anticoagulant;		
KM	anticonvulsant; noctropic; enzyme; DME-10.		
XX			
XX	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	1..50		
FT	Peptide	/label= Signal_peptide	
FT	Domain	32..56	
FT	Protein	/note= "Transmembrane domain"	
FT	Domain	51..642	
FT	Domain	/note= "Mature human DME-10"	
FT	Domain	113..135	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	204..220	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	234..250	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	287..314	
FT	Domain	/note= "Transmembrane domain"	
FT	Domain	463..491	
FT	Domain	/note= "Transmembrane domain"	
XX			
XX	WO200246426-A2.		
XX	13-JUN-2002.		
XX			
XX	04-DEC-2001; 2001WO-US047429.		
XX			
XX	08-DEC-2000; 2000US-0254308P.		
XX	15-DEC-2000; 2000US-0256189P.		
XX	21-DEC-2000; 2000US-0257713P.		
XX	19-JAN-2001; 2001US-0262706P.		
XX	02-FEB-2001; 2001US-0266020P.		
XX			
XX	(INCYTE GENOMICS INC.		
XX			
XX	Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,		
XX	Lee EA, Ding L, Hafalila AJ, Tang YT, Yue H, Tribouley CM, Lu DAM,		
XX	lai PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y;		

Query Match	Best Local Similarity	Score	DB	Length	Matches	548;	Conservative	2;	Mismatches	1;	Indels	0;	Gaps	0
QY	31	GSASGSPQNTLGMTIGQKQVTVLSSPVVNVFLGVPPAAPPLGSLRFTNPQPASPMDNL	90											
DB	92	GSASGSPQNTLGMTIGQKQVTVLSSPVVNVFLGVPPAAPPLGSLRFTNPQPASPMDNL	151											
QY	91	REATSYPLNCIQNSWMLLDQMLKXHPKFCVSEDCYLXNIVYAPAHADTQSKLPVLWF	150											
DB	152	REATSYPLNCIQNSWMLLDQMLKXHPKFCVSEDCYLXNIVYAPAHADTQSKLPVLWF	211											
QY	151	PGAKTKTSASIFDGSALAAVEDVLVVVYVYRLGIFGFPTTMDQHPAGNMAFKQVALS	210											
DB	212	PGAKTKTSASIFDGSALAAVEDVLVVVYVYRLGIFGFPTTMDQHPAGNMAFKQVALS	271											
QY	211	WVQKNIIEFFGDDPSVITFGSAGASISVSLILSPMAKGLPHKAIMESGVALITYLEMD	270											
DB	272	WVQKNIIEFFGDDPSVITFGSAGASISVSLILSPMAKGLPHKAIMESGVALITYLEMD	331											
QY	271	YEKSEDLQVAFHCNNASDSEALLRCLRTKSPKELLTISQTKTSFTRVVDGAFPPNEPL	330											
DB	332	YEKSEDLQVAFHCNNASDSEALLRCLRTKSPKELLTISQTKTSFTRVVDGAFPPNEPL	391											
QY	331	DLISQKAFKALPSILIGVNNHCGGLPMKAPETILSSNKSALHLIONTLHIPPQYLHL	390											
DB	392	DLISQKAFKALPSILIGVNNHCGGLPMKAPETILSSNKSALHLIONTLHIPPQYLHL	451											
QY	391	VANVEYFPHDKSLTETIRSLDLIDGDFVFPVVALITAYTHDAGAPVYFYEFRHRPOCFED	450											
DB	452	VANVEYFPHDKSLTETIRSLDLIDGDFVFPVVALITAYTHDAGAPVYFYEFRHRPOCFED	511											
QY	451	TKPAPVADADADAEFRVFGGAFLLKGLIVTFEGATEEBEKLISRMKMKYMAFPARTGNNGN	510											
DB	512	TKPAPVADADADAEFRVFGGAFLLKGLIVTFEGATEEBEKLISRMKMKYMAFPARTGNNGN	571											
QY	511	DLISLMPVYNTTEQVLTQDLNMSIGORLKEPRVDPFMTSTPLIISASNMVLSPLSTETG	570											

Db 572 DFLTPAVNLTQOYLQDLNLSLGRLEKPEFVFTSTPILLSASDMLHSPSLSTFLS 631

QY 571 LLQPPFFFCAP 581

Db 632 LLQPPFFFCAP 642

RESULT 5

ABM84114
ID ABM84114 standard; protein; 642 AA.

XX AC ABM84114;

XX DT 18-NOV-2004 (first entry)

XX DB Human diagnostic and therapeutic pprotein SEQ ID NO:4363.

XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dthp.

XX OS Homo sapiens.

XX FN WO2004023973-A2.

XX PD 25-MAR-2004.

XX PE 12-SEP-2003; 2003WO-US028227.

XX PR 12-SEP-2002; 2002US-0410259P.

XX PR 12-SEP-2002; 2002US-0410260P.

XX PA (INCY-) INCYTE CORP.

XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F, Hartschorn TA, Suchotolski MT, Altus CM, Pille SJ, Elder LV, Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP, Stevens KM, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH, Percalle CH, Anderson SP, Rioux P, Shen EJ, Wu MC, Stuve LL, Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitek UA, Kirtson ES, Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D, Patuty S, Shi X, Suarez CJ;

XX DR MPI; 2004-329368/30.

XX DR N-PSDB; ACN42766.

XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.

XX PS Claim 27; Page: 190pp; English.

XX CC The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dthp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dthp protein of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX SQ Sequence 642 AA;

Query Match 94.2%; Score 2900; DB 8; Length 642;
Best Local Similarity 99.6%; Pred. No. 7e-262;

Matches 549; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 31 GPSARGPQNRRLIGWIOGKQVTVLGSPPVNVVFLGVPAPAPLGLRFTNPQAPSPMNL 90

Db 92 GPSARGPQNRRLIGWIOGKQVTVLGSPPVNVVFLGVPAPAPLGLRFTNPQAPSPMNL 151

QY 91 REATSYPNLCLQNSWMLLDQMLKVHYPKGVSEDCYLNIYAPAHADTGSKLPLVWF 150

Db 152 REATSYPNLCLQNSWMLLDQMLKVHYPKGVSEDCYLNIYAPAHADTGSKLPLVWF 211

QY 151 PGAPFKTGSASIFDGSALAAVBDVLVYVVOYRLGIFGFTTWDQAPGNAFKQVAAALS 210

Db 212 PGAPFKTGSASIFDGSALAAVBDVLVYVVOYRLGIFGFTTWDQAPGNAFKQVAAALS 271

QY 211 WQKNIEFFGGDPSVTLFGESAGASVSSLLSPMAGLPHKAIMESGVALIPLYEARD 270

Db 272 WQKNIEFFGGDPSVTLFGESAGASVSSLLSPMAGLPHKAIMESGVALIPLYEARD 331

QY 271 YEKSEDLQVNAHFCGNMNSDSEALRLCRLTPSKELTLTSGKTSFTRVVDGAFPPNEPL 330

Db 332 YEKSEDLQVNAHFCGNMNSDSEALRLCRLTPSKELTLTSGKTSFTRVVDGAFPPNEPL 391

QY 331 DLSQKAFKALPSITGVNNHECGFLPMKEAPEIISGNSKSLAHLIIONLIHIPQYHL 390

Db 392 DLSQKAFKALPSITGVNNHECGFLPMKEAPEIISGNSKSLAHLIIONLIHIPQYHL 451

QY 391 VANEYFHDKSLTETRDSDLDDGVFVVPALITARYHRAAGAPVYFEFPHRQCFED 450

Db 452 VANEYFHDKSLTETRDSDLDDGVFVVPALITARYHRAAGAPVYFEFPHRQCFED 511

QY 451 TKPAFVKADHAEVAFVFGAFLKGDIVMEGATEEELSLRKKMKWATFARTGNPQN 510

Db 512 TKPAFVKADHAEVAFVFGAFLKGDIVMEGATEEELSLRKKMKWATFARTGNPQN 571

QY 511 DLSLMPAVNLTQOYLQDLNLSLGRLEKPEFVFTSTPILLSASDMLHSPSLSTFLS 570

Db 572 DLSLMPAVNLTQOYLQDLNLSLGRLEKPEFVFTSTPILLSASDMLHSPSLSTFLS 631

QY 571 LLQPPFFFCAP 581

Db 632 LLQPPFFFCAP 642

RESULT 6

ABP61005
ID ABP61005 standard; protein; 581 AA.

AC ABP61005;

XX DT 10-SEP-2002 (first entry)

XX DE Novel human protein. SEQ ID 92.

XX KW Human; cytostatic; vulnery; antiarteriosclerotic; antiParkinsonian; nootropic; neuroprotective; immunosuppressive; haemostatic; antiinflammatory; cardiant; antilicer; vincinide; antithyroid; cerebroprotective; anorectic; metabolic; vaccine; cancer; infection; wound healing disorders; atherosclerosis; Parkinson's disease; Alzheimer's disease; autoimmune disorder; haematopoietic disorder; inflammation; neoplastic disease; nervous system disorder; cardiovascular disorders; pancreatitis; respiratory disorder; hyperproliferation; systemic autoimmune disease; hyper-immunity; hyperlipidemia; gastrointestinal ulceration; neuropathy; haematological disease; metabolic disease; sperm dysfunction; thyroid disorder; hypothyroidism; brain damage; colitis; cone photo- transduction deficiency; neurological disease; stroke; angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart; trachea; thymus; lymph node; muscular system; obesity; anorexia; growth abnormality; precocious puberty.

XX KW Homo sapiens.

XX OS WO200250105-A1.

XX 27-JUN-2002.
 PD 17-DEC-2001; 2001WO-US049232.
 XX
 PF 19-DEC-2000; 2000US-0256710P.
 XX 20-DEC-2000; 2000US-0257048P.
 PR 09-JAN-2001; 2001US-0260482P.
 PR 30-JAN-2001; 2001US-0264922P.
 PR 06-FEB-2001; 2001US-0266797P.
 PR 19-MAR-2001; 2001US-0276988P.
 PR 04-APR-2001; 2001US-0281535P.
 PR 08-MAY-2001; 2001US-0289622P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
 PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
 XX WPI; 2002-508784/54.
 DR N-PADB; ABQ86170.
 XX
 PT Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX
 PS Claim 1(a); Page 313-314; 335pp; English.
 XX
 CC The invention relates to an isolated polypeptide with signal sequences
 CC which allow it to be secreted extracellularly or membrane associated. The
 CC activity of polypeptides of the invention may be described as,
 CC cytosolic, vulnery, antitumor, antiparkinsonian, neurotrophic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiac, anticancer, antiviral, antihypertensive, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder,
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC developmental abnormalities, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain diseases, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, angiodysplasia,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention
 CC
 CC Sequence 581 AA;
 XX
 SQ
 Query Match 84.7%; Score 2607.5; DB 5; Length 581;
 Best Local Similarity 91.4%; Pred. No. 1.4e-234;
 Matches 499; Conservative 7; Mismatches 25; Indels 15; Gaps 3;
 QY 12 WCFFLIQPLIGHQMGKTPSAEGPQNRRLGNIQSGQVTVGSPVNVVFLGVPPAP 71
 DB 14 WAIWVLAAP-----TKGPSAEGPQNRRLGNIQSGQVTVGSPVNVVFLGVPPAP 65
 QY 72 PLGSLRFTNPQAPSPMNLREATGYPNLCLONSWMILLDDHMLKMHYPRKGVSEDCLYLN 131
 DB 66 PLGSLRFTNPQAPSPMNLREATGYPNLCLONSWMILLDDHMLKMHYPRKGVSEDCLYLN 125
 QY 132 IYAPAHADTGSKLPVLPVFPFGAFKTSASIFDGSALAAEDVLLVVVQYRLGIFGFFTT 191
 DB 126 IYAPAHADTGSKLPVLPVFPFGAFKTSASIFDGSALAAEDVLLVVVQYRLGIFGFFTT 185
 QY 192 WDHAGNNAFKQVVALSVQKNIEFGDPSVITFGSAGAISVSSILISPMKGLF 251

DB 186 WDHAGNNAFKQVVALSVQKNIEFGDPSVITFGSAGAISVSSILISPMKGLF 245
 QY 252 HKAIMESGVALIPYLBAHDYKESBDLOVVAHFCGNNADESLRLCLTKSGKELLTISQ 311
 DB 246 HKAIMESGVALIPYLBAHDYKESBDLOVVAHFCGNNADESLRLCLTKSGKELLTISQ 305
 QY 312 KTKSFTRVVDGAFPNNEPDLISQAFKAPISIIIGVNNHEGFLPMKE--APETLSGSN 369
 DB 306 KTKSFTRVVDGAFPNNEPDLISQAFKAPISIIIGVNNHEGFLPMKE--APETLSGSN 365
 QY 370 KSLAL-----HLIQNLIIPQYHLVANEYFHDHGSITETIRSDLLDLGVFPVVALI 424
 DB 366 KDAALASTAGHFRHGHQHPQYHLVANEYFHDHGSITETIRSDLLDLGVFPVVALI 425
 QY 425 TARYHRDAGAVYFYEFPHRPOCFEDTKPAFVKADHDEVAFVFGAFPLGDIYMEFGAT 484
 DB 426 TARYHRDAGAVYFYEFPHRPOCFEDTKPAFVKADHDEVAFVFGAFPLGDIYMEFGAT 485
 QY 485 EEEKILSRKMMKWTATPARGNPNNDLSLMPAYNLTEOYLQDLNNSLQRLKEPRVD 544
 DB 486 EEEKILSRKMMKWTATPARGNPNNDLSLMPAYNLTEOYLQDLNNSLQRLKEPRVD 545
 QY 545 WTSTRIP 550
 DB 546 WVTGYP 551

RESULT 7
 ADA54807
 ID ADA54807 standard; protein; 525 AA.
 XX
 AC ADA54807;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human protein, SEQ ID 2375.
 XX
 KM Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Neurotrophic;
 KM Gene therapy; human; secretory protein; membrane proteins; cancer;
 KM inflammatory disease; osteoporosis; neurological disease.
 OS Homo sapiens.
 XX
 PN EPI293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-0006586.
 PR 14-SEP-2001; 2001UP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) RES ASSOC BIOTECHNOLOGY.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maunoh Y;
 XX WPI; 2003-395539/38.
 DR N-PADB; ADA53168.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 PS Claim 14; SEQ ID NO 2375; 205pp; English.
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

```

XX Sequence 525 AA;
SQ
Query Match      84.6%; Score 2606; DB 6; Length 525;
      Local Similarity 90.9%; Pred. No. 1.7e-234;
      Matches 501; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

QY 31 GPSAEGPQNRRLGWIQGVQTVLGSPPVNVVFLGVPAPAPPLGSLRFTNPQAPAPMNL 90
DB 25 GPSAEGPQNRRLGWIQGVQTVLGSPPVNVVFLGVPAPAPPLGSLRFTNPQAPAPMNL 84
QY 91 REATSYPLCLQNSEWMLLDQMLKVHYPKFVSESDCLYINITYADAHADTGSKLPVLVWF 150
DB 85 REATSYPLCLQNSEWMLLDQMLKVHYPKFVSESDCLYINITYADAHADTGSKLPVLVWF 144
QY 151 PGAFKTSASIFPDGSALAAVEDVLYVVVQVRLGIFGFFTTMDQAPGNMAFKQVVALS 210
DB 145 PGAFKTSASIFPDGSALAAVEDVLYVVVQVRLGIFGFFTTMDQAPGNMAFKQVVALS 204
QY 211 WYQKNIEPFGDPPSSVTIFGSAGAISVSLLISPMAGLPHKAIMESGVAIIPYLEAHD 270
DB 205 WYQKNIEPFGDPPSSVTIFGSAGAISVSLLISPMAGLPHKAIMESGVAIIPYLEAHD 264
QY 271 YEKSEDLQVNAHFCGNMNSDEBALRCLRTKPSKELTLISQKYSFTRVVDGAFPPNEPL 330
DB 265 YEKSEDLQVNAHFCGNMNSDEBALRCLRTKPSKELTLISQKYSFTRVVDGAFPPNEPL 324
QY 331 DLISQKAFKAIPSIIGVNNHECGFLPMKEAPEIISGNSKSLAHLIONIHLIIPQYVHL 390
DB 325 DLISQKAFKAIPSIIGVNNHECGFLPMKEAPEIISGNSKSLAHLIONIHLIIPQYVHL 384
QY 391 VANEFYDHKSLTEIRDSLLDLGPFVFPVLPALITARVHRAGAVFYEFBRHRQCED 450
DB 385 VANEFYDHKSLTEIRDSLLDLGPFVFPVLPALITARVHRAGAVFYEFBRHRQCED 424
QY 451 TKPAPVNAHDAEVRFPVGAFLKGDIVMFGATEEKEKLSRKMKKYATFARTGNPKN 510
DB 425 -----EGATEEKEKLSRKMKKYATFARTGNPKN 454
QY 511 DLISMPAYNLTEQYQLDLNMSLQRLKEPRVDFMTITPILASDMLHSLSLTFLS 570
DB 455 DLISMPAYNLTEQYQLDLNMSLQRLKEPRVDFMTITPILASDMLHSLSLTFLS 514
QY 571 LLOPFFFCAP 581
DB 515 LLOPFFFCAP 525

RESULT 8
ABUS4639 standard; protein, 581 AA.
ABUS4639;
AC
XX
DT 03-JUN-2003 (first entry)
XX
DB Human NOXV polypeptide #98.
XX
KW Human; NOXV; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS Homo sapiens.
XX
PN MO200281498-A2.
XX
PD 17-OCT-2002.
XX

```

```

PF 03-APR-2002; 2002WO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281116P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.

(CDRA-) CYPAGEN CORP.
PA
XX
PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI Paturajan M, Liu X, Gusev VY, Li L, Vernet CM, Zehruseen BD;
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
PI Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W;
PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI MacDougall JR, Rothenberg MB, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
DR WPI. 2003-046858/04.
DR N-PSDB; ABX72267.
XX
PT New isolated NOXV polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 303; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOXV, and the
XX polynucleotides encoding them. The polypeptides and polynucleotides are
XX useful for diagnosing disease, and screening for potential therapeutic
XX agents. The sequences are useful for treating metabolic disorders,
XX cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
XX stenosis, atrial septal defect (ASD), atrioventricular canal defect,
XX ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
XX septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
XX atherosclerosis, obesity, infectious diseases, anorexia, neurodegenerative
XX disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
XX haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
XX and cancer. Sequences ABUS4542-ABUS4647 represent human NOXV polypeptides
XX of the invention
XX

```

```

SQ      Sequence 581 AA:
Query Match      84.4%; Score 2597.5; DB 6; Length 581;
Best Local Similarity 94.3%; Pred. No. 1.2e-233;
Matches 497; Conservative 4; Mismatches 17; Indels 9; Gaps 3;

QY      31 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPAAPPLGSLRFTNPOQASPMWDL 90
DB      27 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPAAPPLGSLRFTNPOQASPMWDL 86
QY      91 REATSYPNLCLQNSEWLLDQHMLKVHYPKFVSEDCLYINTIYAPAHADTGSKLPLYVWF 150
DB      87 REATSYPNLCLQNSEWLLDQHMLKVHYPKFVSEDCLYINTIYAPAHADTGSKLPLYVWF 146
QY      151 PGGAFTGASATPDSGALAAEDVYVWVVOYRLGTIGFTTWDQAPGMAFKDOYVALS 210
DB      147 PGGAFTGASATPDSGALAAEDVYVWVVOYRLGTIGFTTWDQAPGMAFKDOYVALS 206
QY      211 WVOKNIEFGDPSSVTIFGESAGAISVSLLSPMAKGLFHKAIMESGVALIIPYLEAH 270
DB      207 WVOKNIEFGDPSSVTIFGESAGAISVSLLSPMAKGLFHKAIMESGVALIIPYLEAH 266
QY      271 YKSEBDQVVAHFCGNMADSEALRLCLRTKPSKELLTSQTKSFTRVVDGAFPPNEBL 330
DB      267 YKSEBDQVVAHFCGNMADSEALRLCLRTKPSKELLTSQTKSFTRVVDGAFPPNEBL 326
QY      331 DLSQKAFKAIPIISIGVNNHECGFLPMKE--APEILSGSNKSLAL----HLIQNIHLI 383
DB      327 DLSQKAFKAIPIISIGVNNHECGFLPMRIAAHTATISNDALASTAGHFHRR--HI 384
QY      384 PPOYLHLVANEFYHDKSLTEIRDSLIDLDGDFVFPVPLITARYHRDAGAVYEFERH 443
DB      385 PPOYLHLVANEFYHDKSLTEIRDSLIDLDGDFVFPVPLITARYHRDAGAVYEFERH 444
QY      444 RPOCFEDTPKAFVKADHDEVAFVFGAFLKGDIVMEGATEEELSRKMKKTATATAR 503
DB      445 RPOCFEDTPKAFVKADHDEVAFVFGAFLKGDIVMEGATEEELSRKMKKTATATAR 504
QY      504 TGNPNGNDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFWTSITP 550
DB      505 TGNPNGNDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFWTSITP 551

RESULT 9
ADRI19663
ID      ADRI19663 standard; protein; 618 AA.
XX      AC      ADRI19663;
XX      DT      07-OCT-2004 (first entry)
XX      DE      Human drug metabolizing enzyme (DME)-2 protein sequence.
XX      KW      drug metabolizing enzyme; DME; cytostatic; immunosuppressive;
KW      antiinflammatory; endocrine; ophthalmologic; gastrointestinal;
KW      hepatocellular cancer; cell proliferative disorder; autoimmune disorder;
KW      inflammatory disorder; endocrine disorder; eye disorder;
KW      gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
KW      human.
XX      OS      Homo sapiens.
XX      PN      WO200226988-A2.
XX      PD      04-APR-2002.
XX      PF      28-SEP-2001; 2001MO-US030662.
XX      PR      29-SEP-2000; 2000US-0236947P.
XX      PR      06-OCT-2000; 2000US-0238864P.
XX      PR      20-OCT-2000; 2000US-0242333P.
XX      PR      09-NOV-2000; 2000US-0247581P.
XX      PR      16-NOV-2000; 2000US-0249519P.

PR      22-NOV-2000; 2000US-0252834P.
PR      30-NOV-2000; 2000US-0250567P.
XX      PA      (INCYTE GENOMICS INC.
XX      PI      Azinza Y, Baughin MR, Borowsky ML, Ding L, Dugan BM,
PI      Biliote VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA;
PI      Lai P, Lee BA, Lu DAM, Nguyen DB, Arvizu C, Policky JL, Ramkumar J;
PI      Ring HZ, Santanwala MS, Tang YF, Tribouley CM, Nardinder WK,
PI      Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
XX      DR      MPI: 2002-362498/39.
XX      DR      N-PSDB; ADRI19681.
XX      PT      Human drug metabolizing enzymes, useful in the diagnosis and treatment of
PT      disorders associated with aberrant (DME) activity, e.g., cancer and
PT      autoimmune disorders.
XX      PS      Claim 1; SEQ ID NO 2; 142pp; English.
XX      CC      This invention relates to novel drug metabolizing enzymes (DME) and the
CC      nucleotide sequences which encode them. The invention may be useful for
CC      the development of compounds with a cytostatic, immunosuppressive,
CC      antiinflammatory, endocrine, ophthalmological, gastrointestinal or
CC      hepatocellular activity acting as an agonist or antagonist of drug
CC      metabolizing enzyme activity. The invention may be used in the diagnosis
CC      and treatment of disorders associated with decreased or increased
CC      expression or activity of drug metabolizing enzymes. Such disorders
CC      include cancer, cell proliferative disorders, autoimmune/inflammatory,
CC      endocrine, eye, gastrointestinal (including liver disorders) and
CC      metabolic disorders. The present sequence is that of a human drug
CC      metabolizing enzyme (DME) of the invention. Note: This sequence did not
CC      form part of the printed specification but was obtained in electronic
CC      format from EPO.
XX      SQ      Sequence 618 AA:

Query Match      84.1%; Score 2588.5; DB 5; Length 618;
Best Local Similarity 94.4%; Pred. No. 9.6e-233;
Matches 491; Conservative 1; Mismatches 5; Indels 23; Gaps 1;

QY      31 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPAAPPLGSLRFTNPOQASPMWDL 90
DB      92 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPAAPPLGSLRFTNPOQASPMWDL 151
QY      91 REATSYPNLCLQNSEWLLDQHMLKVHYPKFVSEDCLYINTIYAPAHADTGSKLPLYVWF 150
DB      152 REATSYPNLCLQNSEWLLDQHMLKVHYPKFVSEDCLYINTIYAPAHADTGSKLPLYVWF 211
QY      151 PGGAFTGASATPDSGALAAEDVYVWVVOYRLGTIGFTTWDQAPGMAFKDOYVALS 210
DB      212 PGGAFTGASATPDSGALAAEDVYVWVVOYRLGTIGFTTWDQAPGMAFKDOYVALS 271
QY      211 WVOKNIEFGDPSSVTIFGESAGAISVSLLSPMAKGLFHKAIMESGVALIIPYLEAH 270
DB      272 WVOKNIEFGDPSSVTIFGESAGAISVSLLSPMAKGLFHKAIMESGVALIIPYLEAH 331
QY      271 YKSEBDQVVAHFCGNMADSEALRLCLRTKPSKELLTSQTKSFTRVVDGAFPPNEBL 330
DB      332 YKSEBDQVVAHFCGNMADSEALRLCLRTKPSKELLTSQTKSFTRVVDGAFPPNEBL 391
QY      331 DLSQKAFKAIPIISIGVNNHECGFLPMKEAPEILSGSNKSLALHLIQNIHLI 390
DB      392 DLSQKAFKAIPIISIGVNNHECGFLPM-----HIPOYLHL 428
QY      391 VANEFYHDKSLTEIRDSLIDLDGDFVFPVPLITARYHRDAGAVYEFERH 450
DB      429 VANEFYHDKSLTEIRDSLIDLDGDFVFPVPLITARYHRDAGAVYEFERH 488
QY      451 TKPAFVKADHDEVAFVFGAFLKGDIVMEGATEEELSRKMKKTATATAR 510
DB      489 TKPAFVKADHDEVAFVFGAFLKGDIVMEGATEEELSRKMKKTATATAR 548

```

QY 511 DLSLMPAYNLTEQYLQDLNMSLGGRLKEPRVDFTWTSTIP 550
 DB 549 DLSLMPAYNLTEQYLQDLNMSLGGRLKEPRVDFTWTSTIP 588

RESULT 10

ID ADB64065 standard; protein; 469 AA.

AC ADB64065;

DT 04-DEC-2003 (first entry)

DE Human protein encoded by clone BRAMH20021910.

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 XX cell regeneration; membrane protein; signal transduction-related protein;
 XX transcription-related protein; osteoporosis; neurological disease;
 XX cancer; tumour.

OS Homo sapiens.

PN EP1308459-A2.

XX 07-MAY-2003.

PD 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

PR 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 P1 Yamamoto Y, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 P1 Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maeuho Y;

XX WPI: 2003-450961/43.

DR N-PSDB; ADB62095.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.

XX Claim 1; Page: 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or
 CC peptide of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

XX Sequence 469 AA;

Query Match 79.7%; Score 2454; DB 7; Length 469;
 Best Local Similarity 99.6%; Pred. No. 2.4e-220;
 Matches 467; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 113 MLKVHYPKFVSEDCLYINIVAPAHADTGSKLPLVWFPGGAFKTSASIFDGSALAAVE 172
 DB 1 MLKVHYPKFVSEDCLYINIVAPAHADTGSKLPLVWFPGGAFKTSASIFDGSALAAVE 60
 QY 173 DVLVVVQYRGLGIGFFFTTMDQHPGNMAFPDQVAAISWQKNIIEFGGDSSTVTFEES 232
 DB 61 DVLVVVQYRGLGIGFFFTTMDQHPGNMAFPDQVAAISWQKNIIEFGGDSSTVTFEES 120
 QY 233 AGAIVSSSLISPMAGKLFHAKIMESGVAIIPYLEAHYKESBDLQVVAHFCGNNAASDSE 292
 DB 121 AGAIVSSSLISPMAGKLFHAKIMESGVAIIPYLEAHYKESBDLQVVAHFCGNNAASDSE 180
 QY 293 ALLRCLRTKPSKELLTTSQKTSFTRVVDGAFPPNEPDLISQKAFKAIPSIIGVNNHEC 352
 DB 181 ALLRCLRTKPSKELLTTSQKTSFTRVVDGAFPPNEPDLISQKAFKAIPSIIGVNNHEC 240
 QY 353 GFLPMPKEAPRTLSGSNSLALHLIQLNTLHPQYLHVANEYFHDKSLTEIRPSLDL 412
 DB 241 GFLPMPKEAPRTLSGSNSLALHLIQLNTLHPQYLHVANEYFHDKSLTEIRPSLDL 300
 QY 413 LGDVFVVVALITARYHRDAGAPVYFYEFRHRPOCFEDTKPAFYKADHAEVRFVGGAF 472
 DB 301 LGDVFVVVALITARYHRDAGAPVYFYEFRHRPOCFEDTKPAFYKADHAEVRFVGGAF 360
 QY 473 LKGDIVMEGATEEBEKLSSRQMKYMATFARTGNPNNGDLSLMPAYNLTEQYLQDLNMS 532
 DB 361 LKGDIVMEGATEEBEKLSSRQMKYMATFARTGNPNNGDLSLMPAYNLTEQYLQDLNMS 420
 QY 533 LGGRLKEPRVDFTWTSTIPLLISASDMLHSPSSULTPLSLQPPFFFCAP 581
 DB 421 LGGRLKEPRVDFTWTSTIPLLISASDMLHSPSSULTPLSLQPPFFFCAP 469

RESULT 11

ID ADF50145 standard; protein; 542 AA.

AC ADF50145;

DT 12-FEB-2004 (first entry)

DE Cat cauxin protein SEQ ID NO:2.

XX cat; cauxin; cat kidney disease marker; kidney disease.

OS Felis catus.

PN JP2003250575-A.

PD 09-SEP-2003.

PF 04-MAR-2002; 2002JP-00057908.

PR 04-MAR-2002; 2002JP-00057908.

PA (TOHO-) TOHOKU TECHNOARCH KK.

DR WPI: 2004-002277/01.

DR N-PSDB; ADF50144.

XX Novel cauxin protein or its salt, useful as a cat kidney disease marker,
 PT and for diagnosing cat kidney disease.

XX Claim 3; SEQ ID NO 2; 33pp; Japanese.

XX The present sequence represents a cat cauxin protein (I) or its salt,
 CC which is cat kidney disease marker. Also described: (1) a partial peptide
 CC (II) of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector

CC (IV) comprising (III); (4) a transformed host (V) comprising (III) or
 CC (IV); (5) producing (I) or (II) by culturing (V); (6) an antibody (VI)
 CC which couples specifically with (I) or (II); (7) diagnosing cat kidney
 CC disease which involves measuring (I) quantitatively, and where reduction
 CC of amount of (I) indicates presence of the disease; (8) a cat kidney
 CC disease diagnostic agent comprising (I) labelling agent, a reagent which
 CC measures the biological activity of urinary (I) or (VI); and (9) a cauxin
 CC detection kit which measures cauxin in a test sample. (I) is useful as a
 CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
 CC disease. (I) enables detection of cat kidney disease simply and
 CC correctly. (I) provides an early marker for the disease, and replaces
 CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
 CC blood testing.

CC Sequence 542 AA;

Query Match 64.5%; Score 1985; DB 8; Length 542;
 Best Local Similarity 71.7%; Pred. No. 2.4e-176;
 Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;

QY 31 GPSASGPOHNTLGMVIOGKQVTVLGSPPVNVFLGVFPAPLGSRTNPQAPMNDL 90
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 22 GPADAPVSTRILGWKRGKQTVLGSVPVNMFLGIPYAPPLGRLFRKQPPALPGNDF 81
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 91 REATSYPNLCLONSEMLLDQHLKVHPKFGVSEDCLYLNTIYAPAHADTGSKLPLVWF 150
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 82 RNATSYPKLCFQDLKMLVSYQHVLKVRYPKLEASDCYLNTIYAPAHADNNGSNLVWVWF 141
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 151 PGAFKTSASIFDSSALAAEDVLVVVQYRLGIFGFTTWDQAPGNMAFKDQVAALS 210
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 142 PGAFKMSASISFDSALAAEDVLVTTQYRLGIFGFTDDEHARGMALLDQVALT 201
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 211 WYQKNIFFGDPSSVTTFGESAGAISVSLLISPMAGLFFKAIMESGVALIIPYL-EAH 269
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 202 WVRDNIFFGDPSTVTTFGESAGAISVSLLISPMAGLFFKAIMESGVALIIPLMRPP 261
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 270 DYKSESDIQVAHFGGNASDEALLRCLRTKPSKELLTSOKTGSFTRVVDGAFPPNP 329
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 262 GDERKKDQVLAIRICGCHASDSALLOCLRAKPSSELMDISKULFSPVIDDFPPDP 321
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 330 LDLSQKAFKAIPTIIGVNNHECGFLPMKEAPEILSGSNKSLALHLQNTLHPTQYLH 389
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 322 VALLTQKAFNSVPSIIGVNNHECAFL-STEPSEILGGSNRSALYLVTFTLNIPTQYLH 380
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 390 LVANEYFHDKSLTEIRDSLLDLGDFVFPVVALITARYHRDAGAPVYFYEFRHRPOCFE 449
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 381 LVADHYFYNKHS PVEIRDSFLDLGDFVFPVGVVTARYHRDAGAPVYFYEFRHRPOCLN 440
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 450 DTRPAFYKADHADVRPVGGAFLKGDIVMEFGATEEBSKLSRKMKYKWAFTPARTGNPG 509
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 441 DTRPAFYKADHSDIRPVGGAFLKGDIVMEFGATEEBSKLSRKMKYKWAFTPARTGNPG 500
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 510 NDLSLMPAYNLTEQYLQDLNMSLGQRLKEPRVDFMTSTI 549
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 501 EGVPLWPAVYTGSEQYLKLDLSVSGQLTKQGEVFFMNTI 540
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

RESULT 12

ADFS0147 ADFS0147 standard; protein; 542 AA.

AC ADFS0147;

DT 12-FBB-2004 (first entry)

DE Cat cauxin protein SEQ ID NO:4.

KM cat; cauxin; cat kidney disease marker; kidney disease.

OS Felis catus.

XX JP2003250575-A.

PD 09-SEP-2003.
 XX
 CC 04-MAR-2002; 2002JP-00057908.
 CC
 PF 04-MAR-2002; 2002JP-00057908.
 XX
 PR (TOHO-) TOHOKU TECHNOARCH KK.
 XX
 PA WPI; 2004-002277/01.
 XX
 DR N-PSDB; ADFS0146.
 XX
 PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
 PT and for diagnosing cat kidney disease.
 XX
 PS Disclosure; SEQ ID NO 4; 33pp; Japanese.

The present sequence represents a cat cauxin protein (I) or its salt,
 CC which is cat kidney disease marker. Also described: (1) a partial peptide
 CC (II) of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector
 CC (IV) comprising (III); (4) a transformed host (V) comprising (III) or
 CC (IV); (5) producing (I) or (II) by culturing (V); (6) an antibody (VI)
 CC which couples specifically with (I) or (II); (7) diagnosing cat kidney
 CC disease which involves measuring (I) quantitatively, and where reduction
 CC of amount of (I) indicates presence of the disease; (8) a cat kidney
 CC disease diagnostic agent comprising (I) labelling agent, a reagent which
 CC measures the biological activity of urinary (I) or (VI); and (9) a cauxin
 CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
 CC disease. (I) enables detection of cat kidney disease simply and
 CC correctly. (I) provides an early marker for the disease, and replaces
 CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
 CC blood testing.

XX Sequence 542 AA;

Query Match 64.5%; Score 1985; DB 8; Length 542;
 Best Local Similarity 71.7%; Pred. No. 2.4e-176;
 Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;

QY 31 GPSASGPOHNTLGMVIOGKQVTVLGSPPVNVFLGVFPAPLGSRTNPQAPMNDL 90
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 22 GPADAPVSTRILGWKRGKQTVLGSVPVNMFLGIPYAPPLGRLFRKQPPALPGNDF 81
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 91 REATSYPNLCLONSEMLLDQHLKVHPKFGVSEDCLYLNTIYAPAHADTGSKLPLVWF 150
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 82 RNATSYPKLCFQDLKMLVSYQHVLKVRYPKLEASDCYLNTIYAPAHADNNGSNLVWVWF 141
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 151 PGAFKTSASIFDSSALAAEDVLVVVQYRLGIFGFTTWDQAPGNMAFKDQVAALS 210
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 142 PGAFKMSASISFDSALAAEDVLVTTQYRLGIFGFTDDEHARGMALLDQVALT 201
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 211 WYQKNIFFGDPSSVTTFGESAGAISVSLLISPMAGLFFKAIMESGVALIIPYL-EAH 269
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 202 WVRDNIFFGDPSTVTTFGESAGAISVSLLISPMAGLFFKAIMESGVALIIPLMRPP 261
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 270 DYKSESDIQVAHFGGNASDEALLRCLRTKPSKELLTSOKTGSFTRVVDGAFPPNP 329
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 262 GDERKKDQVLAIRICGCHASDSALLOCLRAKPSSELMDISKULFSPVIDDFPPDP 321
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 330 LDLSQKAFKAIPTIIGVNNHECGFLPMKEAPEILSGSNKSLALHLQNTLHPTQYLH 389
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 322 VALLTQKAFNSVPSIIGVNNHECAFL-STEPSEILGGSNRSALYLVTFTLNIPTQYLH 380
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 390 LVANEYFHDKSLTEIRDSLLDLGDFVFPVVALITARYHRDAGAPVYFYEFRHRPOCFE 449
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 381 LVADHYFYNKHS PVEIRDSFLDLGDFVFPVGVVTARYHRDAGAPVYFYEFRHRPOCLN 440
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 450 DTRPAFYKADHADVRPVGGAFLKGDIVMEFGATEEBSKLSRKMKYKWAFTPARTGNPG 509
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 441 DTRPAFYKADHSDIRPVGGAFLKGDIVMEFGATEEBSKLSRKMKYKWAFTPARTGNPG 500
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
 QY 510 NDLSLMPAYNLTEQYLQDLNMSLGQRLKEPRVDFMTSTI 549
 DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

DB 501 EGVPLMPAYTQSEQYLKILDSVSGVKLKEQGEVEFMNTI 540

RESULT 13

ID ABP61006 standard; protein, 356 AA.

XX ABP61006;

AC ABP61006;

XX 10-SEP-2002 (first entry)

DT 10-SEP-2002 (first entry)

XX Novel human protein. SEQ ID 93.

XX Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian; neurotropic; neuroprotective; immunosuppressive; haemostatic; antinflammatory; cardiant; antidiabetic; vitruicide; antithyroid; cerebroprotective; anorectic; metabolic; vaccine; cancer; infection; wound healing disorders; atherosclerosis; Parkinson's disease; Alzheimer's disease; autoimmune disorder; haematopoietic disorder; inflammation; neoplastic disease; nervous system disorder; cardiovascular disorders; pancreatitis; respiratory disorder; hyperproliferation; systemic autoimmune disease; hyper-immunity; developmental abnormality; gastrointestinal ulceration; neuropathy; haematological disease; metabolic disease; sperm dysfunction; thyroid disorder; hypothyroidism; brain damage; colitis; cone photo-transduction deficiency; neurological disease; stroke; angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart; trachea; thymus; lymph node; muscular system; obesity; anorexia; growth abnormality; precocious puberty.

XX Homo sapiens.

OS WO200250105-A1.

PN 27-JUN-2002.

XX 17-DEC-2001; 2001WO-US049232.

PF 19-DEC-2000; 2000US-0256710P.

XX 20-DEC-2000; 2000US-0257048P.

PR 09-JAN-2001; 2001US-0260482P.

PR 30-JAN-2001; 2001US-0264922P.

PR 06-FEB-2001; 2001US-0266797P.

PR 19-FEB-2001; 2001US-0276988P.

PR 04-APR-2001; 2001US-0281535P.

PR 08-MAY-2001; 2001US-0289622P.

XX (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

PA (GLAX) GLAXO GROUP LTD.

XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KP, Lai Y, Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;

XX MPI: 2002-508784/54.

XX N-PSDB: AEO86171.

DR Secreted proteins and polynucleotides useful as vaccines for preventing or treating various diseases e.g. cancer, wounds, atherosclerosis, Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.

XX Claim 1(a); Page 315; 335pp; English.

XX The invention relates to an isolated polypeptide with signal sequences which allow it to be secreted extracellularly or membrane associated. The activity of polypeptides of the invention may be described as, CC cyostatic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotropic, neuroprotective, immunosuppressive, haemostatic, antiinflammatory, cardiant, antidiabetic, antithyroid, cerebroprotective, anorectic, and metabolic. Polypeptides and polynucleotides of the invention are useful in the treatment, or as a vaccine in the prevention of, cancer, wound healing disorders, infection, atherosclerosis, Parkinson's disease and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,

CC inflammation, neoplastic diseases, nervous system related disorders and CC cardiovascular disorders, pancreatitis, respiratory disorder.

CC hyperproliferation, systemic autoimmune disease, hyper-immunity,

CC developmental abnormality, gastrointestinal ulceration, neuropathy,

CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid disorders e.g. hypothyroidism, brain damages, colitis, cone photo-transduction deficiency, neurological diseases, stroke, angiogenesis, ovulation disorders, diseases in the spinal cord, thyroid gland, heart, trachea, thymus, lymph node and muscular system, obesity, anorexia, growth abnormalities, and alleviation of precocious puberty. The CC sequences given in records ABP60965-ABP61019 represent novel human CC proteins of the invention

XX Sequence 356 AA;

SO Query Match 57.8%; Score 1780; DB 5; Length 356; Best Local Similarity 100.0%; Pred. No. 2e-157; Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ILSPAKGLFKHAIMESGVATIIPTLEAHDYKSEDLQVVAHPCGNASDSEALRCLRTK 301

DB 17 ILSPAKGLFKHAIMESGVATIIPTLEAHDYKSEDLQVVAHPCGNASDSEALRCLRTK 76

QY 302 PSKELLTISQKTSFTRVVDGAFEPNEPDLISQKAFKAIPIIIVNNHECGFLPMKEA 361

DB 77 PSKELLTISQKTSFTRVVDGAFEPNEPDLISQKAFKAIPIIIVNNHECGFLPMKEA 136

QY 362 PEILSGNSKSLALHILQNIHIIPOYLHLVANEYFHDHSLTEIRDSLLDLIGDVFVVP 421

DB 137 PEILSGNSKSLALHILQNIHIIPOYLHLVANEYFHDHSLTEIRDSLLDLIGDVFVVP 196

QY 422 ALITARYRDAGAPVFEFPHRPOCFEDTKPAFKADHAEVRFEGAFIKGDIWVE 481

DB 197 ALITARYRDAGAPVFEFPHRPOCFEDTKPAFKADHAEVRFEGAFIKGDIWVE 256

QY 482 GATEBEKILSRKMKMYATFARTGNPNNDLSLWPAYNLTEQYQLDLMNSIGQRIKEPR 541

DB 257 GATEBEKILSRKMKMYATFARTGNPNNDLSLWPAYNLTEQYQLDLMNSIGQRIKEPR 316

QY 542 VDFWTSITPLILSASDMHSPSSITPLISLIQPFPPFCAP 581

DB 317 VDFWTSITPLILSASDMHSPSSITPLISLIQPFPPFCAP 356

RESULT 14

ID AAE20911 standard; protein, 561 AA.

XX AAE20911;

XX AAE20911;

XX 01-JUL-2002 (first entry)

DT 01-JUL-2002 (first entry)

XX Rat carboxylesterase-like enzyme protein.

DE Rat carboxylesterase-like enzyme.

XX Rat; carboxylesterase-like enzyme; organophosphorus intoxication; osteopathic; gene therapy; osteoporosis; antisense therapy; cytostatic; detoxifying agent; Pager's disease; bone implant degradation; cancer; dental implant; enzyme; gene expression.

XX Rattus sp.

OS WO200206454-A2.

PN 24-JAN-2002.

PD 10-JUL-2001; 2001WO-EP007919.

PF 17-JUL-2000; 2000US-0218564P.

PR (FARB) BAYER AG.

PA Xiao Y;

XX


```

Db      16  GODSASPIRTHHTGVLGSLVHVKGANAGVOFLGI PFAKPPLGRLRAPPEPPESWSGV 75
Qy      91  REATSYPMLCTON-----SEWILLDOHMLKXHYPFKGVSEDCLYNITYAPAHADTGSCLP 145
Db      76  RDGTTHPAMCLODLTAVSEBFL---SQFNMTFPSDSMEDCLYLSITYTPAHSHSGSNLP 131
Qy      146 VLWMPGGAFTGASIFDGSALAYEDVLVVVOYRLGIFGFPTTMDQHA PGNNAFKDQ 205
Db      132 VMMVHGGALVFGMASTYDGSMLALENVVVIIOYRLGLGFFSTGDKATGMGYLDQ 191
Qy      206 VVALSWVQKNIEPFGDPSVTITFGESAGALSVSLLSPMAKGLFHKAIMESGVALIPY 265
Db      192 VVALRMVQONIAHFGNPNDRVTITFGESAGTSVSSLVSPISQGLFHGAIMESGVALLP 251
Qy      266 LEADHYEKSEDL--QVVAHFGCANNASDEBALRCILRTKPSKELTLSOKTSGFTRVVDGA 323
Db      252 LIA---SSADVISTVVANLSACDQVDSBALVGCIRGKSKERIILAIKRPFKMIPGVVDGV 307
Qy      324 FFPNBPDLLSQKAFKAIPSIIGVNNHBCGFLP---MKEAPEILSGSNKSLALHLIQN 379
Db      308 FLPHPOELLSADFOVPSTIVGVNNEFGWLIIPKWMRIYDTQKEMDRASOALQKMLT 367
Qy      380 ILHIPPQYIHLVANEYFHDKHSLTEIRDSLDDLGDVFEVVVPALITARYHRDAGAPVYFY 439
Db      368 LLMLPFTFGDILREBYIGDNGDPQTLQAQFOEMWADSMFVILPALQVAHF-QCSRAPVYFY 426
Qy      440 EFRIRPOCFEDTKPAFVADNADEVRFVFGAFLKGDIVMEGATEEBEKLSSRKMMKTYA 499
Db      427 EFOHQPSTWLNKIRPPHMKADHGDDELPPVF-RSFFGANYIKF---TEEBEQLSRKMMKTYA 482
Qy      500 TPARTGNPNNGDLSIWPAYNLTBOYLQOLDIAMSLSGRLKEPRVDFWTSTIP 550
Db      483 NFAANGNPNNGSLPFWPLPDQBEQYLQNLQPAVGRALKAHRLQFWKCALP 533

```

Search completed: June 15, 2005, 09:37:45
 Job time : 92 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 09:37:21 ; Search time 28 Seconds
(without alignments)
1548.968 Million cell updates/sec

Title: US-10-023-515-2
Perfect score: 3079
Sequence: 1 MPQGLTSSASQWCFLLIQP.....PLSSUTFLSLQPPFFCAP 581

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTOS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3079	100.0	581	4	US-10-023-515-2
2	1230	39.9	559	4	US-09-595-682B-28
3	1230	39.9	559	4	US-09-949-016-6426
4	1230	39.9	577	4	US-09-949-016-9670
5	1150.5	37.4	565	4	US-09-595-682B-21
6	1138.5	37.0	566	4	US-09-264-737-2
7	1130.5	36.7	543	4	US-09-595-682B-26
8	1091.5	35.4	584	2	US-08-845-295A-2
9	1091.5	35.4	584	3	US-09-140-933-2
10	1091.5	35.4	584	3	US-09-146-661-2
11	1091.5	35.4	584	3	US-09-150-515-2
12	1034.5	33.6	539	3	US-09-264-737-1
13	942.5	30.6	454	3	US-08-446-100-26
14	942.5	30.6	454	3	US-08-446-100-28
15	942.5	30.6	454	3	US-08-446-100-30
16	942.5	30.6	454	3	US-08-446-100-31
17	940.5	30.5	454	3	US-08-446-100-27
18	940.5	30.5	454	3	US-08-446-100-29
19	871	29.3	574	4	US-10-023-515-4
20	793	25.8	933	4	US-09-949-016-8386
21	786	25.5	823	4	US-09-949-016-6888
22	782	25.4	953	4	US-09-949-016-8387
23	777.5	25.3	843	4	US-09-491-356C-20
24	775	25.2	575	1	US-08-348-920-1
25	774	25.1	823	1	US-09-491-356C-23
26	774	25.1	575	1	US-08-348-920-2
27	772	25.1	848	4	US-09-491-356C-22

28	769	25.0	836	4	US-09-491-356C-21	Sequence 21, Appl
29	765.5	24.9	614	1	US-08-446-100-25	Sequence 25, Appl
30	764.5	24.8	614	3	US-07-732-962A-2	Sequence 2, Appl
31	764.5	24.8	614	2	US-08-370-156-2	Sequence 2, Appl
32	764.5	24.8	614	3	US-08-446-100-19	Sequence 19, Appl
33	764.5	24.8	614	3	US-08-446-100-21	Sequence 21, Appl
34	764.5	24.8	614	3	US-08-814-095-2	Sequence 2, Appl
35	764.5	24.8	614	5	PCT-US92-06106-2	Sequence 2, Appl
36	764.5	24.8	645	4	US-09-949-016-7063	Sequence 7063, Ap
37	764.5	24.8	645	4	US-09-949-016-7064	Sequence 7064, Ap
38	762.5	24.8	614	3	US-08-446-100-20	Sequence 20, Appl
39	761.5	24.7	614	3	US-08-446-100-23	Sequence 23, Appl
40	760.5	24.7	614	3	US-08-446-100-22	Sequence 22, Appl
41	759.5	24.7	600	2	US-08-370-156-4	Sequence 4, Appl
42	759.5	24.7	600	3	US-08-814-095-4	Sequence 4, Appl
43	759.5	24.7	600	3	US-08-975-084-1	Sequence 1, Appl
44	755.5	24.5	617	2	US-08-370-156-6	Sequence 6, Appl
45	755.5	24.5	617	3	US-08-814-095-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1									
US-10-023-515-2									
Sequence 2, Application US/10023515									
Patent No. 6664091									
GENERAL INFORMATION:									
APPLICANT: Curtiss, Rory A. J.									
TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE									
TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF									
FILE REFERENCE: 10448-122001									
CURRENT APPLICATION NUMBER: US/10/023,515									
CURRENT FILING DATE: 2001-12-18									
PRIOR APPLICATION NUMBER: 60/256,369									
PRIOR FILING DATE: 2000-12-18									
PRIOR APPLICATION NUMBER: 60/279,508									
PRIOR FILING DATE: 2001-03-28									
NUMBER OF SEQ ID NOS: 6									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 2									
LENGTH: 581									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-023-515-2									
Query Match									
Best Local Similarity 100.0%; Score 3079; DB 4; Length 581;									
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MPQGLTSSASQWCFLLIQPLIGHRQKGTGPSAEGPQRNRLGWIQKQVTVLGSPPV	60						
DB	1	MPQGLTSSASQWCFLLIQPLIGHRQKGTGPSAEGPQRNRLGWIQKQVTVLGSPPV	60						
QY	61	NVFLGVPRAAPLIGSLRFTNPQASPMQNLREATSPYLCLQNSBWLDDHMLKVHYPK	120						
DB	61	NVFLGVPRAAPLIGSLRFTNPQASPMQNLREATSPYLCLQNSBWLDDHMLKVHYPK	120						
QY	121	FGVSEDCILYNIYAPAHADTGSKLPLVLPFGCAFKTGSASIFDGSALAAVEDVLVVVVQ	180						
DB	121	FGVSEDCILYNIYAPAHADTGSKLPLVLPFGCAFKTGSASIFDGSALAAVEDVLVVVVQ	180						
QY	181	YRLGIFGFTTWDDQAPNMAFKQVVALSVQKNIIEFGDPPSSVTIFGSAGASIVSS	240						
DB	181	YRLGIFGFTTWDDQAPNMAFKQVVALSVQKNIIEFGDPPSSVTIFGSAGASIVSS	240						
QY	241	LILSPMAGLFRHKIMESGVALIIPYLEADYKESDQVVAHFGNMNSDBEALLRCURT	300						
DB	241	LILSPMAGLFRHKIMESGVALIIPYLEADYKESDQVVAHFGNMNSDBEALLRCURT	300						
QY	301	KPSKELTLTSLQKTSFTRVVVGAFPPNEPLDLSQAKFAIPSIIGVNNHBCGFLLPNKE	360						
DB	301	KPSKELTLTSLQKTSFTRVVVGAFPPNEPLDLSQAKFAIPSIIGVNNHBCGFLLPNKE	360						

Db 301 KPSKELLTISQKTSFTRVVDGAFPPNEPDLDSQAKFAPISIIIVANNHECGFLPMKE 360
Qy 361 APELISGNSKSLAHLIIONILHIPOYLHVANNYFHDKSLTEIRDSLLDLGDVFFV 420
Db 361 APELISGNSKSLAHLIIONILHIPOYLHVANNYFHDKSLTEIRDSLLDLGDVFFV 420
Qy 421 PALITAYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADAEVRFAFGAFKGIYMF 480
Db 421 PALITAYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADAEVRFAFGAFKGIYMF 480
Qy 481 EGATEEKLISRKMMKMTATFARTGNNGDLSIMPAYNLTQYLOLDLMSIGORLKEP 540
Db 481 EGATEEKLISRKMMKMTATFARTGNNGDLSIMPAYNLTQYLOLDLMSIGORLKEP 540
Qy 541 RVDFTSTTPLISASDMLHSPLSTPLSLQPPFFCAP 581
Db 541 RVDFTSTTPLISASDMLHSPLSTPLSLQPPFFCAP 581

RESULT 2
US-09-595-682B-28
/ Sequence 28, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danke, Mary K.
/ APPLICANT: Potter, Philip M.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR FILING DATE: 2000-01-16
/ PRIOR APPLICATION NUMBER: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR APPLICATION NUMBER: PCT/US99/03171
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 28
/ LENGTH: 559
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-595-682B-28

Query Match 39.9%; Score 1230; DB 4; Length 559;
Best Local Similarity 46.7%; Pred. No. 4.8e-125;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

Qy 31 GPSAEGPQNRTRLGIWIOGQVTVLGSPPVAVNVPFLGVPFAAPPLGSLRFTNQPASPMYDL 90
Db 26 GQDSASPIRTHHTGQVLSLVHVGANAGVOTFLGIPRAKPLGLRPAPEPPESWGV 85
Qy 91 REATSYPNLCION-----SEWLLDQHLKVHYKFGVSEDCLYLNTIYAPAHADTGSKL 145
Db 86 RDGTHPAMCLODLAVESSEFL-----SQFNMTPSDMSSEDCLYLSTYTPAHSHSGNLP 141
Qy 146 VLWVPGAFKTSASIFDGSALAAVEDVLYVVOYRLGTFGFTTMDQAPGMNAFKDQ 205
Db 142 VMMVHIGGALVFGMAISLVDSMLALSNVAVVYIYRLGVAGFSTGDKATGNMGYLDQ 201
Qy 206 VAALSWQKNIIEFGDPSVTTFEGESAGAISVSLLISPMAGLFFKAIMESGVAILIPY 265
Db 202 VAALRWQONIAHFGNDPRTVTFEGESAGTSVSLVSPISQGLFHGAIMESGVAILIPG 261
Qy 266 LEAHDYKESDL--QVVAHFCGNNAASEALRLCLRTKPSKELLTISQKTSFTRVVDGA 323
Db 262 LIA-----SSADVISTVAVNLASACQVDSEALVGLCRGSKKEILAIKPKMIPGVVDG 317
Qy 324 FFPNEPDLDSQAKFAPISIIIVANNHECGFLP-----MKEAPEIISGNSKSLAHLIION 379
Db 318 FLPRHPQELIASADFOQVPSIVGVNNNEFGWLLPKWRIYDTQKEMDRASQALQKMLT 377
Qy 380 ILHIPOYLHVANNYFHDKSLTEIRDSLLDLGDVFFVVALITARYHRDAGAPVYF 439

Db 378 LMLPPTFGDLIREEYIGDNGDPQTLQAQFQEMADSMFVLPALQVANH-QCSAPVYFY 436
Qy 440 EFRHRPQCFEDTKPAFVADAEVRFAFGAFKGIYMEGATEEKLISRKMMKMTA 499
Db 437 EFQHPQSLKNIIRPPHMKADHGDLELPVYF-RSFGQNYIKF---TEEBEQLSRKMMKMTA 492
Qy 500 TFARTGNNGDLSIMPAYNLTQYLOLDLMSIGORLKEPVDFTSTP 550
Db 493 NFARNGNNGSGLPHPWLPFDQEBQYLOLNTQPAVGRALKAKHRLQPKKALP 543

RESULT 3
US-09-949-016-6426
/ Sequence 6426, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6426
/ LENGTH: 559
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-6426

Query Match 39.9%; Score 1230; DB 4; Length 559;
Best Local Similarity 46.7%; Pred. No. 4.8e-125;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

Qy 31 GPSAEGPQNRTRLGIWIOGQVTVLGSPPVAVNVPFLGVPFAAPPLGSLRFTNQPASPMYDL 90
Db 26 GQDSASPIRTHHTGQVLSLVHVGANAGVOTFLGIPRAKPLGLRPAPEPPESWGV 85
Qy 91 REATSYPNLCION-----SEWLLDQHLKVHYKFGVSEDCLYLNTIYAPAHADTGSKL 145
Db 86 RDGTHPAMCLODLAVESSEFL-----SQFNMTPSDMSSEDCLYLSTYTPAHSHSGNLP 141
Qy 146 VLWVPGAFKTSASIFDGSALAAVEDVLYVVOYRLGTFGFTTMDQAPGMNAFKDQ 205
Db 142 VMMVHIGGALVFGMAISLVDSMLALSNVAVVYIYRLGVAGFSTGDKATGNMGYLDQ 201
Qy 206 VAALSWQKNIIEFGDPSVTTFEGESAGAISVSLLISPMAGLFFKAIMESGVAILIPY 265
Db 202 VAALRWQONIAHFGNDPRTVTFEGESAGTSVSLVSPISQGLFHGAIMESGVAILIPG 261
Qy 266 LEAHDYKESDL--QVVAHFCGNNAASEALRLCLRTKPSKELLTISQKTSFTRVVDGA 323
Db 262 LIA-----SSADVISTVAVNLASACQVDSEALVGLCRGSKKEILAIKPKMIPGVVDG 317
Qy 324 FFPNEPDLDSQAKFAPISIIIVANNHECGFLP-----MKEAPEIISGNSKSLAHLIION 379
Db 318 FLPRHPQELIASADFOQVPSIVGVNNNEFGWLLPKWRIYDTQKEMDRASQALQKMLT 377
Qy 380 ILHIPOYLHVANNYFHDKSLTEIRDSLLDLGDVFFVVALITARYHRDAGAPVYF 439
Db 378 LMLPPTFGDLIREEYIGDNGDPQTLQAQFQEMADSMFVLPALQVANH-QCSAPVYFY 436
Qy 440 EFRHRPQCFEDTKPAFVADAEVRFAFGAFKGIYMEGATEEKLISRKMMKMTA 499
Db 437 EFQHPQSLKNIIRPPHMKADHGDLELPVYF-RSFGQNYIKF---TEEBEQLSRKMMKMTA 492

Qy 500 TFAATGPNNGNDLSMPAYNLTEOYLQDLNMSLGRLEKPRVDFTSTIP 550
Db 493 NFAANGPNNGGLPHWPLFDEBEYLOLNLQPAVGRALKARLQFWKALP 543

RESULT 4
US-09-949-016-9670
; Sequence 9670, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9670
; LENGTH: 577
; TYPE: PRF
; ORGANISM: Human
US-09-949-016-9670

Query Match 39.9%; Score 1230; DB 4; Length 577;
Best Local Similarity 46.7%; Pred. No. 5.1e-125;
Matches 244; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

Qy 31 GPSAGBQFORTRLGKIOGKQVTVLGSPPVNVFLGVPAPPLGSLRPTNQPASPMNDL 90
Db 44 GDBASPIRTHHTQVGLSVHVKAGANAGVQTFIGIPRAKPLGRLRPAPEBPESMSGV 103
Qy 91 REASTYNNLCION-----SEWLLDDQMLKHYKPKGVSEDCLYLNTIYAPHAATGSLCP 145
Db 104 RDGTTTHPAMCLODTJAVESEPL-----SOFNMTFPSDSSEDCLYLSTYTPHSHGSLNP 159
Qy 146 VLWHPGGAFTGTSASIFDGSALAAYEDVLVVVQYRLGFGFTTMDQHPGMMAFMDQ 205
Db 160 VMWVHIGGALVFGNASTLYDSMLALLENVVVITIQRLGVGFSTGDKATGMWYLDQ 219
Qy 206 VAALSWYQKNIIEFGDPSSTYIFGESAGALSVSLLSPMAKGLFKKAIMESGVAIIPY 265
Db 220 VVALRWQONIAHGGNPDRTTIFGESAGTSVSLVSPISQGLFGALMESGVALLPG 279
Qy 266 LEANDYKESDL--QVVAHFCGNASDSEALRLCLRTKPSKELLTTSOKTSPTRVYDGA 323
Db 280 LIA-----SSADVISTVANLSACDOVSEALVGLRGSKEEIIAINKPFKPIRGVVDGV 335
Qy 324 FFPNEPLDLSQKAFKAIPISTIIGVNHGCGFLP-----MKEAPEILSGSNKSLALHLION 379
Db 336 FLPHPEBLASADFOVPVPSIVGVNNEFGWLIKRWKRIYDTQKEMDRBSAOLLOKMLT 395
Qy 380 ILHLPOLYHLVANEYFHDKSLTEIRDSLDLIDGVFEVVPALITARYHRDAGAPVYFY 439
Db 396 LLMPLPFPGDLAREYIGDNGDPQTLAQOPEMMADSWFVIRALQVAF--QCSAPVYFY 454
Qy 440 EFRHRPQCBEYTKAFYKADADAVRPVFGGAFKGLDIWMEGATBEKLSRRQMKYWA 499
Db 455 EFOHQPSMLKIRIPPHKADHDGDELPIFYF--RSFPGGVYIKF---TEEBEOLSRQMKYWA 510
Qy 500 TFAATGPNNGNDLSMPAYNLTEOYLQDLNMSLGRLEKPRVDFTSTIP 550
Db 511 NFAANGPNNGGLPHWPLFDEBEYLOLNLQPAVGRALKARLQFWKALP 561

RESULT 5
US-09-595-682B-21

; Sequence 21, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ. ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 565
; TYPE: PRF
; ORGANISM: Oryctolagus cuniculus
US-09-595-682B-21

Query Match 37.4%; Score 1150.5; DB 4; Length 565;
Best Local Similarity 43.9%; Pred. No. 2.5e-116;
Matches 252; Conservative 75; Mismatches 180; Indels 67; Gaps 13;

Qy 12 WCFPLLIQPLGRHOKRTPSAGFORTRLGKIOGKQVTVLGSPPVNVFLGVPAPAP 71
Db 2 WLCAALASLAACTAWGH--PSAP--PVVDVHGVKGLKFSLBGFAQPAVAVFLGVPAPK 58
Qy 72 PLGSLRTPPOPASPMNDLREASTYNNLCIONSEWLLDDQML-----KVHYP--KFGV 123
Db 59 PLGSLRTPPOPASPMNDLREASTYNNLCIONSEWLLDDQML-----KVHYP--KFGV 112
Qy 124 SEDCLYNTIYAPHAADTGSRLPYLVWHPGGAFTGTSASIFDGSALAAYEDVLVVVQYRL 183
Db 113 SEDCLYNTIYAPHAADTGSRLPYLVWHPGGAFTGTSASIFDGSALAAYEDVLVVVQYRL 172
Qy 184 GIRGFTTMDQHPGMMAFMDQVVALSWYQKNIIEFGDPSSTYIFGESAGALSVSLLIL 243
Db 173 GIMGFSTGDEHSGNMGHLDQVALLRWVDNTANFGSDGGSYITIFGESAGGSVSILL 232
Qy 244 SPMAKGLFKKAIMESGVAIIPYLEANDYKESDL--LQVVAHFCGNASDSEALRLCLRTK 301
Db 233 SPLTKNLFHRAISBSGVALLSSL-----FRKQVTSIAEKIIEAGCKTTTSAWVHCLRQK 288
Qy 302 PSKELL--TISOK-----TKSFTRVYDGAFFPNEPLDLSQKAFKAIPISTIG 346
Db 289 TEBELMEVTLKMFPMALDVGDPREKNTAFLTVYIDGVLLPKAPAEILAEKKYNMLPYWVG 348
Qy 347 VNHGCGFLP-----KEAPEILSGSNKSLALHLIONILHLPOLYHLVA 392
Db 349 INQOEFMIIPMOWGLGYLSEGLDKQATATELWKS-----YPIVNVSKELTPVAT 399
Qy 393 NEYFHDGSLTEIRDSLDLIDGVFEVVPALITARYHRDAGAPVYFEPHRRPQCBEYTK 452
Db 400 EKYTGTDTPVKKDLFLDMLADLFGVPSVNVARHHRDGAFLMEYIRFSPSSDMK 459
Qy 453 PAFYKADADAVRPVFGGAFKGLDIWMEGATBEKLSRRQMKYWAFTATGPNNGNDL 512
Db 460 PKTYIDGHGDEIPSVLGAPFLK-----EGATEBEIRLTSQWVKYANFARNGPNNGGL 513
Qy 513 SLMPAYNLTEOYLQDLNMSLGRLEKPRVDFTSTIP 546
Db 514 POWPAYDYKKGYLQIGATTQAAOKLKDKEVAFWT 547

RESULT 6
US-09-264-737-2
; Sequence 2, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:

```

; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; TITLE OF INVENTION: Expression of Esterase Enzymes
; FILE REFERENCE: 38-21(10551) RUS3 Pyridine Tolerance
; CURRENT APPLICATION NUMBER: US/09/264,737A
; EARLIER FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; EARLIER FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Rabbit
; US-09-264-737-2

```

Query Match 37.0%; Score 1138.5; DB 3; Length 566;
Best Local Similarity 43.6%; Pred. No. 5.1e-115;

Matches 250; Conservative 75; Mismatches 176; Indels 73; Gaps 14;

```

Qy 12 WCFPLIQLPGLHROKGTGPSAEGPQRNRLGWIQKQVTLGSPVNVVFLGVPPAP 71
Db 9 WLF-----LAACIATMGH--PSAP-PVVDIVHGKVLGKFSLEGFAPVAVFLGVPPAP 59
Qy 72 PLGSLRFTNPQAPSPMDNLRKATSYPNLCLQNSEWMLLDQHML-----KHYAP-KFGV 123
Db 60 PLGSLRFAPQAPSPMDNLRKATSYPNLCLQNSEWMLLDQHML-----KHYAP-KFGV 113
Qy 124 SEDCLYNTIYAPAHADTGSKLPLYWFPFGAFKTSASIFDSALAAVEDLVVVVQYRL 183
Db 114 SEDCLYNTIYTPADLTGKGRLLPVMVWIHGGIWMVGASTYDGLALSAHENVVVTIYQYRL 173
Qy 184 GIGFETFTWDQAPGNMAFKDQVALSWVQKNIIEFGGDPSSVTITGESAGAISSVSLIL 243
Db 174 GIMGFSTGDEHSGKNGHLDQVALRWVODIANFGDPSGVTITGESAGAGSISLIL 233
Qy 244 SPMAKGLFHAKMESGVALIPLYEAHDYKSED--LQVVAHFCGNNASDEALRLCLRTK 301
Db 224 SPLTKMLFHRAISEGVALISL-----FRKNTKSLAEKIAIEAGCKTTTSAVMVHCLRQK 289
Qy 302 PSKELL--TLGSK-----TKSFTRVNDGAFPPNEBLLISOKAFKAIPISTIG 346
Db 290 TEEBLMEVTLKMKFMALDVGDPKENTAFLLTVIVGVLIPKAPABITLAEKKYKMLPYVWG 349
Qy 347 VNNHCEGFLPM-----KEAPILSGSNKSLALHLIQLNLIHPQYLHLVA 392
Db 350 INQOEFGWIIIPMOMGLYPLSEKLDQKATATLWKS-----YPIVNVSKELTPVAT 400
Qy 393 NEYFHDKSLTEIRDSLLDLGDVFFVVPALITARYHRDAGAPVYFEFRHRPQCFEDTK 452
Db 401 EKYIGGTDPPVKKDLFLDMLADLLFGVSVVVAHRHDAGAPTYMEYRVRPSPSSDMR 460
Qy 453 PAFVKADHADEVRFVFGAFLKGDIVMEFEGATEEEKLSRKMKYKMAFPARTGNNGNDL 512
Db 461 PKTVIGDHDEIFSVLGAFLK-----EGATEEIRKLSKVMKYKMAFPARNGNNGGL 514
Qy 513 SLMPAYNLTEOYLQDLNMSLQRLKEPRVFWT 546
Db 515 POWPAYDYKEGYLQIGATTQAAGQLKDXEVAFWT 548

```

```

RESULT 7
; US-09-595-682B-26
; Sequence 26, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SC-0005

```

```

; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 26
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; US-09-595-682B-26

```

Query Match 36.7%; Score 1130.5; DB 4; Length 543;
Best Local Similarity 43.7%; Pred. No. 3.5e-114;
Matches 249; Conservative 75; Mismatches 179; Indels 67; Gaps 13;

```

Qy 12 WCFPLIQLPGLHROKGTGPSAEGPQRNRLGWIQKQVTLGSPVNVVFLGVPPAP 71
Db 2 WLCALALASLAACTAWGH--PSAP-PVVDIVHGKVLGKFSLEGFAPVAVFLGVPPAP 58
Qy 72 PLGSLRFTNPQAPSPMDNLRKATSYPNLCLQNSEWMLLDQHML-----KHYAP-KFGV 123
Db 59 PLGSLRFAPQAPSPMDNLRKATSYPNLCLQNSEWMLLDQHML-----KHYAP-KFGV 112
Qy 124 SEDCLYNTIYAPAHADTGSKLPLYWFPFGAFKTSASIFDSALAAVEDLVVVVQYRL 183
Db 113 SEDCLYNTIYTPADLTGKGRLLPVMVWIHGGIWMVGASTYDGLALSAHENVVVTIYQYRL 172
Qy 184 GIGFETFTWDQAPGNMAFKDQVALSWVQKNIIEFGGDPSSVTITGESAGAISSVSLIL 243
Db 173 GIMGFSTGDEHSGKNGHLDQVALRWVODIANFGDPSGVTITGESAGAGSISLIL 232
Qy 244 SPMAKGLFHAKMESGVALIPLYEAHDYKSED--LQVVAHFCGNNASDEALRLCLRTK 301
Db 223 SPLTKMLFHRAISEGVALISL-----FRKNTKSLAEKIAIEAGCKTTTSAVMVHCLRQK 288
Qy 302 PSKELL--TLGSK-----TKSFTRVNDGAFPPNEBLLISOKAFKAIPISTIG 346
Db 289 TEEBLMEVTLKMKFMALDVGDPKENTAFLLTVIVGVLIPKAPABITLAEKKYKMLPYVWG 348
Qy 347 VNNHCEGFLPM-----KEAPILSGSNKSLALHLIQLNLIHPQYLHLVA 392
Db 349 INQOEFGWIIIPMOMGLYPLSEKLDQKATATLWKS-----YPIVNVSKELTPVAT 399
Qy 393 NEYFHDKSLTEIRDSLLDLGDVFFVVPALITARYHRDAGAPVYFEFRHRPQCFEDTK 452
Db 400 EKYIGGTDPPVKKDLFLDMLADLLFGVSVVVAHRHDAGAPTYMEYRVRPSPSSDMR 459
Qy 453 PAFVKADHADEVRFVFGAFLKGDIVMEFEGATEEEKLSRKMKYKMAFPARTGNNGNDL 512
Db 460 PKTVIGDHDEIFSVLGAFLK-----EGATEEIRKLSKVMKYKMAFPARNGNNGGL 513
Qy 513 SLMPAYNLTEOYLQDLNMSLQRLKEPRVFWT 542
Db 514 POWPAYDYKEGYLQIGATTQAAGQLKDXEVAFWT 543

```

```

RESULT 8
; US-08-845-295A-2
; Sequence 2, Application US/08845295A
; Patent No. 5817490
; GENERAL INFORMATION:
; APPLICANT: Hubbs, John C.
; TITLE OF INVENTION: Enzymatic Process for the Manufacture of
; TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
; TITLE OF INVENTION: 2-Keto-L-Gulonic Acid
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Eastman Chemical Company
; STREET: P. O. Box 511
; CITY: Kingsport

```


STATE: Tennessee
COUNTRY: USA
ZIP: 37662-5075
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,295A
FILING DATE: 25-April-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,879
FILING DATE: 17-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cheryl J. Tubach
REGISTRATION NUMBER: 38,346
REFERENCE/DOCKET NUMBER: 70432
TELECOMMUNICATION INFORMATION:
TELEPHONE: 423-229-6189
TELEFAX: 423-229-1239
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-845-295A-2

Query Match 35.4%; Score 1091.5; DB 2; Length 584;
Best Local Similarity 43.8%; Pred. No. 7,4e-110;

Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;

12 WCFPLLIQPLIGHKQMGKTPSAGPQQRNRLGWIQKQVTV--LGSPVPYNNVFLGVPA 69
2 WLLPLVLITLSSASATW--AGQPASPPVVDTAQGRVLGKRVSLGLELAFTQPAVFLGVPA 59
70 APPLGSLRFTNPOPASPMNDLRBATSYNLCQNS--EWLLD-----QHMLKVHPK 120
60 KPLGSLRFPAPQAPAEPMSPVKNTTSYPMCCQDPVVGQMTSDLEFTNFGKERLTLEF-- 117
121 FGVEDCLYINITYAPAHADTGSKLPLVWPFPGAFTKGSASIFDGSALAAYED--VLVVV 178
118 ---SEDCLYINITYPADLTGKGRPLVWVWIHGGGLVVGAGPMYDGVVLAHENEFTVVVVA 174
179 VQYRLGIFGFFTTWDQAHAPGMAFKDQVVALSWVQKNIFFGGDPSSTITGES--AGAI 236
175 IQYRLGIFGFSTGDEHSRGWGHLDQVVALHWQENIANFGGDGSGVTITFGESFTAGGE 234
237 SVSSILISPMKGLFHKAIMESGVAIIPYLEAHDYKESDQVVAHFCGNNAASDEALLR 296
235 SVSVLVLSPLAKNLFHRAISSGVALTVLVKMKMAAKQAIATVLAGCKTTTSAVFTFVH 294
237 CLRTKPSKELLTISQKTSFT-----RVVDGAFFPNPRLDLSOK--AFK 339
295 CLRKQSEBELDLTLTKKFTLTDHFGRDRESHPFLPTVVDGVLLPKMBEETIAEKDFTFN 354
340 AIPSIIGVNNHECGFLP-----MKBAPEILSGSNKSLAHLIONTLHIPPQ 386
355 TVPYIVGINKQEGFGLPLTWMMGFLPSBGKLDQKNTSLMKS-----YPLANI-----PE 404
387 YLHLVANEYFDKX-----SLTEIRDSLDDLGDFVFNVPALITARKYRDGAPYTFEF 441
405 ELTPVAT--FTDKYLGSTDDPVKKKDLFLDLMGDVVFQPSVTAVARQRHDGAPTYMEF 462
442 RHRPCCFED--TKPAFVADHADDEVRFVGGAFLLGDTVMFPGATBEKTLSSRKMKXTWA 499
463 QYRBSFSSDKTKTKPTVYIGDHGDELFSVFGPLLGND-----APEEVSISKTYMKEMA 516
500 TPAATGNDGNDLSLP--AYNLTQOYLQDLNMSLAGRLKEPRVDFTSTI 549
517 NFARGNGNGSGLPHMFPPTWYDQEGVYLQIGVNTQAARLKGEBVAFVNDLL 568

RESULT 9

US-09-140-933-2

Sequence 2, Application US/09140933

Patent No. 6022719

GENERAL INFORMATION:

APPLICANT: Hubbs, John C.

TITLE OF INVENTION: Enzymatic Process for the Manufacture of

TITLE OF INVENTION: Acarboxic Acid, 2-Keto-L-Gulonic Acid, and Esters of

TITLE OF INVENTION: 2-Keto-L-Gulonic Acid

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Eastman Chemical Company

STREET: P.O. Box 511

CITY: Kingsport

STATE: Tennessee

COUNTRY: USA

ZIP: 37662-5075

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: Microsoft Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/140,933

FILING DATE: 27-August-98

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/017,879; 08/845,295

FILING DATE: 17-May-96; 25-April-97

ATTORNEY/AGENT INFORMATION:

NAME: Cheryl J. Tubach

REGISTRATION NUMBER: 38,346

REFERENCE/DOCKET NUMBER: 70432

TELECOMMUNICATION INFORMATION:

TELEPHONE: 423-229-6189

TELEFAX: 423-229-1239

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 584 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

MOLECULE TYPE: protein

US-09-140-933-2

Query Match 35.4%; Score 1091.5; DB 3; Length 584;
Best Local Similarity 43.8%; Pred. No. 7,4e-110;

Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;

12 WCFPLLIQPLIGHKQMGKTPSAGPQQRNRLGWIQKQVTV--LGSPVPYNNVFLGVPA 69
2 WLLPLVLITLSSASATW--AGQPASPPVVDTAQGRVLGKRVSLGLELAFTQPAVFLGVPA 59
70 APPLGSLRFTNPOPASPMNDLRBATSYNLCQNS--EWLLD-----QHMLKVHPK 120
60 KPLGSLRFPAPQAPAEPMSPVKNTTSYPMCCQDPVVGQMTSDLEFTNFGKERLTLEF-- 117
121 FGVEDCLYINITYAPAHADTGSKLPLVWPFPGAFTKGSASIFDGSALAAYED--VLVVV 178
118 ---SEDCLYINITYPADLTGKGRPLVWVWIHGGGLVVGAGPMYDGVVLAHENEFTVVVVA 174
179 VQYRLGIFGFFTTWDQAHAPGMAFKDQVVALSWVQKNIFFGGDPSSTITGES--AGAI 236
175 IQYRLGIFGFSTGDEHSRGWGHLDQVVALHWQENIANFGGDGSGVTITFGESFTAGGE 234
237 SVSSILISPMKGLFHKAIMESGVAIIPYLEAHDYKESDQVVAHFCGNNAASDEALLR 296
235 SVSVLVLSPLAKNLFHRAISSGVALTVLVKMKMAAKQAIATVLAGCKTTTSAVFTFVH 294
237 CLRTKPSKELLTISQKTSFT-----RVVDGAFFPNPRLDLSOK--AFK 339
295 CLRKQSEBELDLTLTKKFTLTDHFGRDRESHPFLPTVVDGVLLPKMBEETIAEKDFTFN 354

Qy 340 AIPSIIGVNNHCEGLP-----MKAPRILSGSNKSLALHILIONIHIPQ 386
 Db 355 TVPIYIVGINKOEFWMLPPTMMGFPLSGKLDQKTAISLWKS-----YPIANI-----PE 404
 Qy 387 YLHLVANEYFHDK-----SLTEIRDSLIDLGLDVPFVVPALITARYHRDAGAPVYEF 441
 Db 405 ELTPVAT--FTDKYLGSTDDPVKKDLFLDMDGVGVPSVVALQHRDAGAPTYMYEF 462
 Qy 442 RHRQCED--TPAPYKADHADEVRFVFGAFLKGDIVFEGATEBEKLSRMMKYWA 499
 Db 463 QYRFSFSDKFTKPYTIGDHGDEIFSVFGFPLKGD-----APEBEVSLSKTYWKFMA 516
 Qy 500 TFARTGNPNNDLSLWP--AYNLTEOYLQDLNMSLGGORLKEPRVFWMTSTI 549
 Db 517 NFARSGNPNGEGLPHWPFMTYDDEBGLQIGVTOAKRLKGEVAFMNDL 568

RESULT 10
 US-09-146-661-2
 / Sequence 2, Application US/09146661
 / Patent No. 6136575
 / GENERAL INFORMATION:
 / APPLICANT: Hubbs, John C.
 / TITLE OF INVENTION: Enzymatic Process for the Manufacture of
 / TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of 2-Keto-L-Gu
 / NUMBER OF SEQUENCES: 3
 / CORRESPONDENCE ADDRESSES:
 / ADDRESSER: Eastman Chemical Company
 / STREET: P.O. Box 511
 / CITY: Kingsport
 / STATE: Tennessee
 / COUNTRY: USA
 / ZIP: 37662-5075
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 3.5 inch disk
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM: Windows 95
 / SOFTWARE: Microsoft Word
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/146,661
 / FILING DATE: 03-September-98
 / CLASSIFICATION:
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 60/017,879; 08/845,295
 / FILING DATE: 17-May-96; 25-April-97
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Cheryl J. Tubach
 / REGISTRATION NUMBER: 38,346
 / REFERENCE/DOCKET NUMBER: 70432
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 423-229-6189
 / TELEFAX: 423-229-1239
 / INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 584 amino acids
 / TYPE: Amino Acid
 / TOPOLOGY: Linear
 / MOLECULE TYPE: protein
 / US-09-146-661-2

Query Match 35.4%; Score 1091.5; DB 3; Length 584;
 Best Local Similarity 43.8%; Pred. No. 7.4e-110;
 Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;
 Qy 12 WCFPLLIQPLIGHRQMGKTPSAGPQRNTRLGWIOGKQVTV--LGSPPVNVNVLGVPEFA 69
 Db 2 WLLPLVLTSLASSATW--AGQPAAPPVVDTAQGRVGLKTYVSLBGLAFTGQVNAFLGVPEFA 59
 Qy 70 APPFGSRFTNPPQAPSPMDLREATSYRNLCLQNS--EWLLD-----QHMLKHYVPE 120
 Db 60 KPFLGSRFAPPPQAPSPMDLREATSYRNLCLQNS--EWLLD-----QHMLKHYVPE 117

Qy 121 FGVSDECLYNIYAPAHADTGSKLPVLVWPPGAFKTSASIPGSAALAYED--VLVVV 178
 Db 118 ---SEDCIYNIYTPADLTGKGRLPVWVHGGGLVGLGAMVYQGVVLAHAEFTVVVA 174
 Qy 179 VQYRIGIGFTTMDQAPGWAFFQDVALSWQKNIFFGSDPSSVTTFGES--AGAI 236
 Db 175 IQYRIGIGFTTMDQAPGWAFFQDVALSWQKNIFFGSDPSSVTTFGES--AGAI 234
 Qy 237 SVSLSLSPMAKGFPHKAIMESGVALIPYLAHYKESDELQVVAHCCNNAASELLR 296
 Db 235 SVSVALVSLPAKULPHRAISSGVALYVALVKMKAAQIAVLAQCKTTTSVFTVH 294
 Qy 297 CLRTYPSKELLITLSQKTSFT-----RVWDGAPFPNPELIDLQK--APK 339
 Db 295 CLRQKSEDELIDLTLKMKFLLTLDPHGDRSHPELPVVDGVLLPKRPEILAEKQFTFN 354
 Qy 340 AIPSIIGVNNHCEGLP-----MKAPRILSGSNKSLALHILIONIHIPQ 386
 Db 355 TVPIYIVGINKOEFWMLPPTMMGFPLSGKLDQKTAISLWKS-----YPIANI-----PE 404
 Qy 387 YLHLVANEYFHDK-----SLTEIRDSLIDLGLDVPFVVPALITARYHRDAGAPVYEF 441
 Db 405 ELTPVAT--FTDKYLGSTDDPVKKDLFLDMDGVGVPSVVALQHRDAGAPTYMYEF 462
 Qy 442 RHRQCED--TPAPYKADHADEVRFVFGAFLKGDIVFEGATEBEKLSRMMKYWA 499
 Db 463 QYRFSFSDKFTKPYTIGDHGDEIFSVFGFPLKGD-----APEBEVSLSKTYWKFMA 516
 Qy 500 TFARTGNPNNDLSLWP--AYNLTEOYLQDLNMSLGGORLKEPRVFWMTSTI 549
 Db 517 NFARSGNPNGEGLPHWPFMTYDDEBGLQIGVTOAKRLKGEVAFMNDL 568

RESULT 11
 US-09-150-515-2
 / Sequence 2, Application US/09150515
 / Patent No. 6271006
 / GENERAL INFORMATION:
 / APPLICANT: Hubbs, John C.
 / TITLE OF INVENTION: Enzymatic Process for the Manufacture of
 / TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
 / NUMBER OF SEQUENCES: 3
 / CORRESPONDENCE ADDRESSES:
 / ADDRESSER: Eastman Chemical Company
 / STREET: P.O. Box 511
 / CITY: Kingsport
 / STATE: Tennessee
 / COUNTRY: USA
 / ZIP: 37662-5075
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 3.5 inch disk
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM: Windows 95
 / SOFTWARE: Microsoft Word
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/150,515
 / FILING DATE: 09-SEP-1998
 / CLASSIFICATION:
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 60/017,879; 08/845,295
 / FILING DATE: 17-May-96; 25-April-97
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Cheryl J. Tubach
 / REGISTRATION NUMBER: 38,346
 / REFERENCE/DOCKET NUMBER: 70432
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 423-229-6189
 / TELEFAX: 423-229-1239
 / INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 584 amino acids

TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-09-150-515-2

Query Match 35.4%; Score 1091.5; DB 3; Length 584;

Best Local Similarity 43.8%; Pred. No. 7,4e-110;
Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;

```

QY 12 WCFPLLPGLGHRMGKTPSAGBQRTKLGIMIGKQVTV--LGSPVNVVFLGVPFA 69
DB 2 WLPPLVTLSSLSASAT--AGOPASPPVVDYTAQGRVLGKVSLEGLAFTQPVAVFLGVPFA 59
QY 70 APPLSLRPTNPQAPSPMDLREATSPYLCLONS--EWLLLD-----QHLKHYK 120
DB 60 KPLPLSLRPAQAPAPSPMSPVNTTSPYMCODPVEQMTSDLPFNFGKRLTIEF-- 117
QY 121 FGVSBDCLYLNTPAPAHADTGSKLPLYVWFGGAFKTSASIFDSALAAVED--VLVVV 178
DB 118 ---SDCLYLNITYTADLTGRRLPYMWVVIHGGGLVLGGAPMYDGVVLAHNFVTVVVA 174
QY 179 VQYRIGTIGFTTQOHAAGNNAFKDQVAAALSMVQKNIFFGSDPSVTIFGES--AGAI 236
DB 175 IQYRIGTIGFTTQOHAAGNNAFKDQVAAALSMVQKNIFFGSDPSVTIFGESFTAGGE 234
QY 237 SVSLLISPMAGLPHKAIMESGVAIIPYLAHDIKESDQLQVVAHFCGNMNASDSEALLR 296
DB 235 SVSLLISPMAGLPHKAIMESGVAIIPYLAHDIKESDQLQVVAHFCGNMNASDSEALLR 294
QY 297 CLRTKPSKELLTLTQKTKSFT-----RVVDGAFPPNEPLDLSQK--AFK 339
DB 295 CLRTKPSKELLTLTQKTKSFT-----RVVDGAFPPNEPLDLSQK--AFK 339
QY 340 AIPSTIGVNNHCGFLP-----MKAPELLSSGNSKSLAHLIQLNIHHPQ 386
DB 340 AIPSTIGVNNHCGFLP-----MKAPELLSSGNSKSLAHLIQLNIHHPQ 386
QY 355 TVPYIVGINKORFGMLPTMGPFLSSEGLDQKTSLSLMS-----YPIINI-----PE 404
DB 355 TVPYIVGINKORFGMLPTMGPFLSSEGLDQKTSLSLMS-----YPIINI-----PE 404
QY 367 YLHLVANEYFHDKH-----SLTEIRDSLLDLGADVFPVVPALITARYHNDAGAPYFYE 441
DB 405 ELLPVAT--FTDKYVAGTDDPVKKDLFLDLMDGVVFGVSVTVARQHNDAGAPYFYE 462
QY 442 RHPOCFED--TKPAFVKAHDAEVRFYVFGAFLKGDIVMEFEGATEEERLSRKMKKYWA 499
DB 463 QYRPFSSDKFTKPKTVIGDHDEIFSVRGFPLKGD-----APBEVSLSKTMKFWA 516
QY 500 TPARTGNPNGLSLWP--AYNLTEOYLQDLNMSLIGORLKEPRVDFTSTI 549
DB 517 NFARSGNPNGBGLPMWPFMYDQEGYLIQIVNTQAAKRLKGEVAFVNDLL 568

RESULT 12
US-09-264-737-1
; Sequence 1, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
; APPLICANT: Beng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; FILE REFERENCE: 38-21(1051) RLE3 Pyridine Tolerance
; CURRENT FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Rabbit
US-09-264-737-1
Query Match 33.6%; Score 1034.5; DB 3; Length 539;

```

Best Local Similarity 42.2%; Pred. No. 1.1e-103;
Matches 234; Conservative 71; Mismatches 176; Indels 73; Gaps 13;

```

QY 32 PSABGPQNTKRLGIMIGKQVTVLGSPPVNVVFLGVPFAAPLGSIRFTNPQASPMNLR 91
DB 2 PSAP-PVVDYVAGKVLGKVSLEGLAFTQPVAVFLGVPFAAPLGSIRFTNPQASPMNLR 60
QY 92 EATSPNCLQSEWLLDOHML-----KHYF-KGVSEDCYLNTPAPAHADTGSK 143
DB 61 NNTTSIFPMCSSA---VSGHMLSELFTNRKENITLKF--SDCLYLNITYTADLTGR 114
QY 144 LPVLVWFGGAFKTSASIFDSALAAVEDVLVVVQYRIGTIGFTTQOHAAGNNAFK 203
DB 115 LPVWVVIHGGGLMVGASTYDGLALSAHENVVTVIQRIGTIGGFGFNIDE-----L 166
QY 204 DOVALALSMVQKNIFFGSDPSVTIFGESAGALSVSLLISPMAGLPHKAIMESGVAI 263
DB 167 FLVAVNRWVDNIANFGDPGSGVTIFGESAGOSVLSLLESTLNLPHRALISEGVAL 226
QY 264 PYLAHDIKESD--LQVVAHFCGNMNASDSEALLRCLRTKPSKELL--TLQK----- 312
DB 227 SSL-----FRKTKSLAEKIALIAGCKTTSVAMVHCLQKTEBELMEVTLKMKFMA 282
QY 313 -----TKSPTRVVDGAFPPNEPLDLSQKAFKAIPIIGVNNHCGFLP----- 358
DB 283 GDPKENTAFLLTVIDGVLLPKAPAEIYEKKYNNMLPYWGINQGFWMIPWQMLGYPLS 342
QY 359 -----KEAPELLSSGNSKSLAHLIQLNIHHPQYLAHANEYFHDKSLTEIRPSLD 412
DB 343 EGKLDQKATBELMS-----YPIVNSKELLTPATEKYLQGTDDPVKKDLFLDM 393
QY 413 LGDVFVPVVALITARYHNDAGAPYFYEFRHPOCFEDTKPAFVKAHDAEVRFYVFGAF 472
DB 394 LADLLFGVSVVVAHNDAGAPYFYEFRHPOCFEDTKPAFVKAHDAEVRFYVFGAF 453
QY 473 LKGDIVMEFEGATEEERLSRKMKKYWATFATGNPNGLSLWPAYNLTEOYLQDLNMS 532
DB 454 LK-----EGATEEERLSRKMKKYWATFATGNPNGLSLWPAYNLTEOYLQDLNMS 507
QY 533 LGORLKEPRVDFT 546
DB 508 AAKRLKGEVAFVNDLL 521

RESULT 13
US-08-446-100-26
; Sequence 26, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535

```

```

; REFERENCE/DOCKET NUMBER: broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
;
US-08-446-100-26

Query Match      30.6%; Score 942.5; DB 3; Length 454;
Best Local Similarity 44.9%; Pred. No. 9.5e-94;
Matches 200; Conservative 64; Mismatches 150; Indels 31; Gaps 6;

QY 125 EDCLYNTIYAPAHADTGSKLPVLVFPFGGAFKTSASIFDGSALAAVEDLVVVVQYRLG 184
DB 1 EHCLYNTIYTPADLTQKQRLPVMVWIHGGLMVGAASTYDGLAALAHENVVVVTIQYRLG 60
QY 185 IFGFPTTDOAHAPGMAFQDOVAALSWQKNIEFPGDPSVTITFGESAGATSVSSLLIS 244
DB 61 IWGFSTGDHSRGNWGHLDQVAALRWODNIASFGNPGSVTITFGESAGGSVSVALVS 120
QY 245 PMAKGLFHKAIMESGVAIIPYLEAHDYER--SEDLQVVAHFCCGNNASDSEALRLCLRTKP 302
DB 121 PLAKLNFHRAISSEGVALTSLVKKGDVKKPLAEQIAITA--GCKTTTSAAVHCLRQKT 177
QY 303 SKELLTISOK-----TKSFTRVVDGAFPPNEPDLISQAKFAIPSLIGV 347
DB 178 EEBLETTTLKIGNSYLTWYRETQRESTLLGTVIDMLLKTPELQRENFHTVPYMWGI 237
QY 348 NNHECGFLPMKEAPEILSGS--NKSIALHLION---ILHIPOYLALVANEYFHDKSL 402
DB 238 NKQEPGMILPMQMSYPLBSGQLODKTMSLISGSIPLPAIAKELIPATEKTLGGTDDT 297
QY 403 TEIRDSLIDLIDGVFVVVPALITARYHRDAGAPVYFEFRHRPOCFEDTKPAFYKADHAD 462
DB 298 VKKCDLILDLIDVWFGVPSVIVARNHRDAGAPTYMFEQYRPSFSOMKPKTVIGDHG 357
QY 463 EVRFVFGAFLKGDIMVPEGATEEKKLSRKMKYATFARTGNPNGNLISLMPAYNLTE 522
DB 358 ELFSVFGAPFLK-----EGASEEBIRLSKMKYKFWANFARNGNPNKGLPHMPEYNOKE 411
QY 523 QYLQLDINMSLGQRLKEPRVDFTWS 547
DB 412 GYLQIGANTQAQKLDKKEVAFWTN 436

RESULT 14
US-08-446-100-28
; Sequence 28, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Okeana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glennia
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEtical: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
;
US-08-446-100-28

Query Match      30.6%; Score 942.5; DB 3; Length 454;
Best Local Similarity 44.9%; Pred. No. 9.5e-94;
Matches 200; Conservative 65; Mismatches 149; Indels 31; Gaps 6;

QY 125 EDCLYNTIYAPAHADTGSKLPVLVFPFGGAFKTSASIFDGSALAAVEDLVVVVQYRLG 184
DB 1 EHCLYNTIYTPADLTQKQRLPVMVWIHGGLMVGAASTYDGLAALAHENVVVVTIQYRLG 60
QY 185 IFGFPTTDOAHAPGMAFQDOVAALSWQKNIEFPGDPSVTITFGESAGATSVSSLLIS 244
DB 61 IWGFSTGDHSRGNWGHLDQVAALRWODNIASFGNPGSVTITFGESAGGSVSVALVS 120
QY 245 PMAKGLFHKAIMESGVAIIPYLEAHDYER--SEDLQVVAHFCCGNNASDSEALRLCLRTKP 302
DB 121 PLAKNLFHRAISSEGVALTSLVKKGDVKKPLAEQIAITA--GCKTTTSAAVHCLRQKT 177
QY 303 SKELLTISOK-----TKSFTRVVDGAFPPNEPDLISQAKFAIPSLIGV 347
DB 178 EEBLETTTLKIGNSYLTWYRETQRESTLLGTVIDMLLKTPELQRENFHTVPYMWGI 237
QY 348 NNHECGFLPMKEAPEILSGS--NKSIALHLION---ILHIPOYLALVANEYFHDKSL 402
DB 238 NKQEPGMILPMQMSYPLBSGQLODKTMSLISGSIPLPAIAKELIPATEKTLGGTDDT 297
QY 403 TEIRDSLIDLIDGVFVVVPALITARYHRDAGAPVYFEFRHRPOCFEDTKPAFYKADHAD 462
DB 298 VKKCDLILDLIDVWFGVPSVIVARNHRDAGAPTYMFEQYRPSFSOMKPKTVIGDHG 357
QY 463 EVRFVFGAFLKGDIMVPEGATEEKKLSRKMKYATFARTGNPNGNLISLMPAYNLTE 522
DB 358 ELFSVFGAPFLK-----EGASEEBIRLSKMKYKFWANFARNGNPNKGLPHMPEYNOKE 411
QY 523 QYLQLDINMSLGQRLKEPRVDFTWS 547
DB 412 GYLQIGANTQAQKLDKKEVAFWTN 436

RESULT 15
US-08-446-100-30
; Sequence 30, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B

```

APPLICANT: Lockridge, Oksana
 TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hendricks and Assoc.
 STREET: 9669 A Main Street, P.O. Box 2509
 CITY: Fairfax
 STATE: VA
 COUNTRY: US
 ZIP: 22031
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,100
 FILING DATE: 19-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Hendricks, Glenna
 REGISTRATION NUMBER: 32,535
 REFERENCE/DOCKET NUMBER: Broomfield
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 425-4250
 TELEFAX: (703) 425-3767
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 454 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ANTI-SENSE: YES
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 ORGANISM: human esterases
 US-08-446-100-30
 Query Match 30.6%; Score 942.5; DB 3; Length 454;
 Best Local Similarity 44.9%; Pred. No. 9.5e-94;
 Matches 200; Conservative 64; Mismatches 150; Indels 31; Gaps 6;
 QY 125 EDCIYINITYAPAHADTSGSKLPVLMWFGGAKRTGSASIFDGSALAAYEDVLYVWVQYRLG 184
 DB 1 EHCYLYINITYPADLTJKNRLPVMWVHGGGLMVGAASTYDGLALAHENVVVVTIQYRLG 60
 QY 185 IFGFFTTMDHAPGMAFKDOVALSWQKNIFFPGDPSVTIFGSAGAI SVSSLTLS 244
 DB 61 IWGFFSTGDEHSRGNWGHLDVVALRWQDNINASHGNSPGSVTIFGSAGBSVSVLVLS 120
 QY 245 PMAKGLPKAIAMESGVAILPYLEADHYEK--SEDLQVVAHFCGNNASDSEALLRCLRTKP 302
 DB 121 PLAKNLIFRAISEGSVALTSVLKKGVKPLAEQAIATF---GCKTTSAAMVHCLRQKT 177
 QY 303 SKELLTTSQK-----TKSFTRVVDGAFPPNEPLDLSQAKAIPSIIGV 347
 DB 178 EBELETTTLKIGNSVLTWYRETORHSTLGTVIDGMLLKTPEBLQERENFTVPYWGVI 237
 QY 348 NNHCEGFLLPKBAPEILSGS--NKSIALHLIION---ILHI PPOYLHVANEYFPHDKSL 402
 DB 238 NKQFPGWILIFQALMSYPLSEGLDOKTAMSLGSPILFALAKELIPEATEKYLGGTDDT 297
 QY 403 TEIRDSLDLGVDYFVVVVALITARYHRDAGAPVYFYEFRHRPOCFEDTKPAFYKADHAD 462
 DB 298 VKKDDLLIDLADVWFGVSVIVARNHRDAGAPTYWVFQYRPSFSSDMKPKTVIGDHGD 357
 QY 463 EVRFVFGGAPFKGDIWMEGATTEEEKLSRKQMKTKWATPARTGNPNNDLSLMPAYNLTE 522
 DB 358 ELFSVFGAPFLK-----EGASEEIEIRLSKVMKFWANFARNGNPNNGKGLPHMPEYNOKE 411
 QY 523 QYLDLNMSLGQRLKEPRVDFTWS 547

DB 412 GYLDIGANTQAQKTKDKREVAFTN 436
 Search completed: June 15, 2005, 09:43:54
 Job time : 30 secs

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 15, 2005, 09:43:26 ; Search time 7228 Seconds

(without alignments)
3894.920 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079

Sequence: 1 MPGLTSSASQWCFLLQP.....PLSLTFLSLQPPFFCAP 581

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/gen2.1/USPTO.epool/6664091/runat_14062005_133640_10461/app.query.fasta.1.775
-DB=genmb1 -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=humand0.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091 @CGN.1.1.4198 @runat_14062005_133640_10461 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3079	100.0	1746	6	AR438765 Sequence
2	3079	100.0	1746	6	AX600126 Sequence
3	3079	100.0	2158	6	AR438764 Sequence
4	3079	100.0	2158	6	AX600124 Sequence

5	2915	94.7	2047	9	BC069501 Homo sapi
6	2900	94.2	2232	6	AX477711 Sequence
7	2881.5	93.6	2298	9	BC039073 Homo sapi
8	2877	93.4	2092	6	AX746724 Sequence
9	2877	93.4	2092	6	AX090927 Sequence
10	2867	92.6	1896	9	BC069548 Homo sapi
11	2606	84.6	1962	6	AX714052 Sequence
12	2606	84.6	1962	6	AX056109 Homo sapi
13	2588.5	84.1	1857	6	AX958392 Sequence
14	2505	81.4	1716	6	CO733044 Sequence
15	2250	73.1	1728	4	AB186392 Canis fam
16	2094.5	68.0	1728	4	AB186393 Mus muscu
17	2026	65.8	2139	4	AB045377 Fells cat
18	1676.5	54.4	1739	4	AB060873 Macaca fa
19	1304.5	42.4	2262	5	DUKTER L05493 Anas platyr
20	1285	41.7	1931	10	BC034178 Mus muscu
21	1285	41.7	1937	10	BC034180 Mus muscu
22	1285	41.7	1952	10	BC034191 Mus muscu
23	1285	41.7	1961	10	BC024552 Mus muscu
24	1285	41.7	1965	10	BC015290 Mus muscu
25	1285	41.7	1974	10	BC031170 Mus muscu
26	1282.5	41.7	1776	10	AB110073 Mus muscu
27	1281	41.6	2012	10	BC055062 Mus muscu
28	1276.5	41.5	1848	10	D50577 Mus muscu
29	1271.5	41.3	1913	10	BC058815 Mus muscu
30	1264.5	41.1	1927	10	AB010632 Rattus no
31	1264.5	41.1	2257	5	BX934916 Gallus ga
32	1263.5	41.0	2295	5	BC074230 Xenopus l
33	1260.5	40.9	2846	10	BC027185 Mus muscu
34	1257.5	40.8	2228	10	D50580 Rattus norv
35	1249.5	40.6	1942	10	BC015286 Mus muscu
36	1239.5	40.3	2083	5	BC064228 Xenopus t
37	1238	40.2	1856	10	AY034877 Rattus no
38	1236	40.1	2249	5	BX934906 Gallus ga
39	1233.5	40.1	1780	10	BC025812 Mus muscu
40	1233.5	40.1	1786	10	BC024082 Mus muscu
41	1233.5	40.1	1786	10	BC031295 Mus muscu
42	1233.5	40.1	1788	10	BC026643 Mus muscu
43	1233.5	40.1	1789	10	BC024491 Mus muscu
44	1233.5	40.1	1790	10	BC024517 Mus muscu
45	1233.5	40.1	1805	10	BC025537 Mus muscu

ALIGNMENTS

RESULT 1
AR438765 AR438765 1746 bp DNA linear PAT 20-FEB-2004
LOCUS Sequence 3 from patent US 6664091.
DEFINITION AR438765
ACCESSION AR438765
VERSION AR438765.1 GI:42663740
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1746)
AUTHORS Curtis,R.A.J. and Silos-Santiago,I.
TITLES 53010, a human carboxylesterase family member and uses thereof
JOURNAL Patent: US 6664091-A 3 16-DEC-2003;
FEATURES
source Location/Qualifiers
1..1746
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3,08e-252 Length: 1746
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-023-515-2 (1-581) x AR438765 (1-1746)

QY 1 MetProGlnIleuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 20
Db 1 ATGCGAAGGAGACTTACTTATCTGCTTCAACATGGGCTTTTCTCGATTCCAGACC 60
QY 21 LeuLeuGlnIleuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 40
Db 61 CTGTGGGAGACAGACAGTGGGAGAAACCTGGGCTTCTGCTAAGGGCCACAGAGAAC 120
QY 41 ThrArgLeuGlnIleuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 60
Db 121 ACCAGGCTGGGATGGATTCAGGGCAAGCATGCTGCTGGGAGACCTGGGCTGG 180
QY 61 AsnValPheLeuGlnIleuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 80
Db 181 AACGTTCTCTGGAGATCCCTTCTGCTGCTCCCGCTGGAGATCCCTGGATTCAGAAC 240
QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgIleuAlaThrSerTrpProAsnLeuCys 100
Db 241 CCGCAGCCTGACATCGCCTGGGATTACTTGGAGAGACCACTCTACCTAATTTGTGC 300
QY 101 LeuGlnAsnSerGlnIleuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 120
Db 301 CTCAGAACTCAGAGTGGCTGCTTGAATCAACACATGCTCAAGGTGATTAACCGAAA 360
QY 121 PheGlnValSerGlnAspCysLeuTrpLeuAsnIleuTrpAlaProAlaHisAlaAspThr 140
Db 361 TTCGAGGTGTCAAGAGACTGGCTTACCTGAACTTAATGCGCTGCCACGCCATACA 420
QY 141 GlnSerIleuLeuProValLeuValIleuThrPheProGlnIleuAlaPheIleuThrIleSerAla 160
Db 421 GGCTCAAGCTCCCGCTTGGTGGTTCACAGAGGTGGCTTCAAGCTGGCTCAGCC 480
QY 161 SerIlePheAspGlnSerAlaLeuAlaIleuThrValIleuValValIleuValGln 180
Db 481 TCCATCTTGAAGGGTCCGCCCTGGCTGCTTGAAGAGAGTGGTGGTGGTGGTGGTGG 540
QY 181 TyrArgLeuGlnIleuPheGlnPhePheThrTrpTrpAspGlnHisAlaProGlnValAsnTrp 200
Db 541 TACCGGCTAGGAATTTGGTCTTCAACACATGGAGTCAAGCATGCCGGGGAACCTGG 600
QY 201 AlaPheIleuAspGlnValAlaAlaLeuSerTrpValGlnIleuValAsnIleuGlnPhePheGln 220
Db 601 GCCTTCAAGAGACAGGTGGCTGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 660
QY 221 GlnAspProSerSerValThrIlePheGlnIleuSerAlaGlnAlaIleSerValSerSer 240
Db 661 GGGGAGCCCACTGTGATCCATCTTGGGAGTCCGGGAGCCATTAAGTCTTCTAGT 720
QY 241 LeuIleLeuSerProMetAlaIleuGlnIleuPheHisIleuValAlaIleMetGlnSerGlnVal 260
Db 721 CTTATACGTGTCCATGGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTG 780
QY 261 AlaIleIleProTrpLeuGlnIleuAlaHisAspTrpGlnIleuSerGlnIleuValVal 280
Db 781 GCCATCATCCCTTACCTGAGAGCCCATGATTAAGAGAGTGGAGACCTGCAAGTGGT 840
QY 281 AlaHisPheCysGlnIleuAsnAlaSerAspSerGlnIleuLeuAsnArgCysLeuArgThr 300
Db 841 GCACATTTCTGTGTAACTAATGCTGAGACTCTGAGGCCCTGAGAGTGGTGGAGACA 900
QY 301 LysProSerIleuGlnIleuLeuThrLeuSerGlnIleuTrpIleuSerPheMetArgValVal 320
Db 901 AAACCTTCAAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
QY 321 AspGlnValaPhePheProAsnGlnProLeuAspLeuSerGlnIleuValaPheVala 340
Db 961 GATGGTCTTCTTCTTCTAATGAGCTCTAGATTAATGCTCTCAAGAAAGCATTTAAAGCA 1020
QY 341 IleProSerIleuIleuGlnIleuValaAsnHisGlnCysGlnIleuLeuProMetIleuGln 360
Db 1021 ATTCTTTCATCATCGAGTCAATTAACACAGAGTGGTGGTGGTGGTGGTGGTGGTGG 1080

QY 361 AlaProGlnIleuLeuSerGlnIleuAsnIleuSerAlaLeuHisIleuGlnIleuVal 380
Db 1081 GCTCTGAGATCTCAGTGGCTCCAAAGATCCCTTGGCTCCATCTGATTCAAAACATC 1140
QY 381 LeuHisIleProProGlnIleuHisIleuValAlaAsnGlnIleuPheHisAspIleuHis 400
Db 1141 CTGACATCCCGCTCAGATTTTGGACCTTGTGGTAAATGATCTTCATGAGAACGAC 1200
QY 401 SerLeuThrGlnIleuArgAspSerLeuLeuAspLeuGlnIleuValaPheValVal 420
Db 1201 TCCCTGATGAAATCCGAGACAGTCTTGGACCTTCTGAGAGATGCTTTTGTGGTC 1260
QY 421 ProAlaLeuIleuTrpAlaArgTrpHisArgAspAlaGlnIleuProValIleuPheTrpGln 440
Db 1261 CCTGACATGATCAAGCTCGATATCAAGAGATCTGGTGCACCTGCTACTTCTATGAG 1320
QY 441 PheArgHisArgProGlnCysPheGlnAspThrIleuSerProAlaPheValIleuValaPheHis 460
Db 1321 TTTGGCAGCCGGCTCAGTGGCTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACAC 1380
QY 461 AlaAspGlnValaArgPheValaPheGlnIleuAlaPheLeuIleuValaMetPhe 480
Db 1381 GCTGATGAAGTCCGCTTGTGTTGCTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1440
QY 481 GlnIleuValaThrGlnIleuGlnIleuLeuSerArgIleuMetIleuTrpAlaThr 500
Db 1441 GAAGAGCCACAGAGAGAGAGAGATTAAGAGCCGAGAAATGATGAATCTGGGCTAAC 1500
QY 501 PheAlaArgThrGlnIleuAsnProAsnGlnIleuAspLeuSerLeuTrpProAlaTrpAsnLeu 520
Db 1501 TTTGCTGAAACCGGAAATCTTAATGGAGACAGCTGCTGAGGCACTTAATCTG 1560
QY 521 ThrGlnIleuTrpLeuGlnIleuAspLeuAsnMetSerLeuGlnIleuArgLeuIleuPro 540
Db 1561 ACTGAGCAGTACCTCCAGCTGAGACTTGAACATGAGCTCGAGACAGAGACTCAAGAACCG 1620
QY 541 ArgValaAspPheThrTrpIleuSerTrpIleuLeuIleuSerAlaSerAspMetLeuHis 560
Db 1621 CGGGTGAATTTTGGACACACACATCCCTTATCTGCTGCTCCAGACATGCTCCAC 1680
QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
Db 1681 AGTCCCTCTTCTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 581 Pro 581
Db 1741 CCT 1743

RESULT 2
AX600126 1746 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 3 from Patent WO0250256.
DEFINITION AX600126
ACCESSION AX600126.1 GI:28400192
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
Curtis, R.A. and Siloe-Santiago, I.
53010, a novel human carboxylesterase family member and uses
thereof
Patent: WO 0250256-A 3 27-JUN-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
location/Qualifiers
FEATURES
source
1..1746
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 3,086-252 Length: 1746
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-023-515-2 (1-581) x AK600126 (1-1746)

```

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuLLeuGlnPro 20
DB 1 ATGCACAGGAGACTTACTTCACTCTTCAACATGAGTGTCTTTCTGATCTCCAGCCC 60
QY 21 LeuLeuGlyHISArgGlnTrpGlyLeuThrGlyProSerAlaGluGlyProGlnArgAsn 40
DB 61 CTGTTGGGACACAGACAGTGGGGAAACCTGGCCCTTCTGCTGAAGGGCCACAGGAGAAC 120
QY 41 ThrArgLeuGlyTrpIleGlnGlyLeuValThrValLeuGlySerProValProVal 60
DB 121 ACCAGGCTGGAGTGGATTCAGGGGCAAGCACTGTGCTGGGAAGCCCTGTGCTGTG 180
QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
DB 181 AACGTTCTCTCGAGAGTCCCTTTGCTGCTCCCGCTGGAGTCCCTGGATTTACGAAC 240
QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGluAlaThrSerTrpProAsnLeuCys 100
DB 241 CCGCAGCCTGCATCCCTCGGGATTAATTGGCAAGAGCACCTCTACCTTAATTTGTGC 300
QY 101 LeuGlnAsnSerGluTrpLeuLeuLeuAspGlnHISMetLeuValValHISTrpProLys 120
DB 301 CTCGAAACTCAGAGTGGCTGCTCTTAATCAACACATCTCAAGGTGATTAACCGGAAA 360
QY 121 PheGlyValSerGluAspCysLeuTrpLeuAsnIleTrpAlaProAlaHISAlaAspThr 140
DB 361 TTCGAGGTGAGAAAGACTGCTCTTACCTGAACATCTATGGCTGCCACGCCCATACA 420
QY 141 GlySerIleLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySerAla 160
DB 421 GGCTCCAAAGCTCCCGCTGTGGTGTGGTCCAGAGAGGTGCTTCAAGCTGGCTCAGCC 480
QY 161 SerIlePheAspGlySerAlaLeuAlaIleTrpGluAspValLeuValValGln 180
DB 481 TCCATCTTTGATGGTCCGCCCTGGCTGCTATGAGAGAGTCTGCTGTGTGCTCCAG 540
QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHISAlaProGlyAsnTrp 200
DB 541 TACGGGCTAGGAATTTGGTTCTTTCACACATGGGATCAGCATGCTCCGGGAACTGG 600
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGluPhePheGly 220
DB 601 GCCTTCAGAGACACAGTGGCTGCTGTCTGCTGGTCCAGAAAGACATCAAGTCTTCGGGT 660
QY 221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerAlaSerSer 240
DB 661 GGGGAGCCCAAGCTCTGTGACATCTTTGGCGAGTCCGGGGGCCATTAAGTGTTCCTGAT 720
QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHISLysAlaIleMetGluSerGlyVal 260
DB 721 CTTATACGTGTCCCATAGGCCAAAGGCTTAATCCAAAGGCATCATGAGAGTGGGGTG 780
QY 261 AlaIleIleProTrpLeuGlnAlaHISAspTrpGlyLysSerGluAspLeuGlnValVal 280
DB 781 GCCATCATCCCTTACCTGAGGCCCATGATTAATGAGAGAGAGAGACCTTCAGAGTGGTT 840
QY 281 AlaHISpHeCysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuArgCysLeuArgThr 300
DB 841 GCACATTTCTGTGGTGAACATGCTCAAGCTCTGAGGCCCTCTGAGGTGCTTGAGAGACA 900
QY 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysTrpLysSerPheThrArgValVal 320
DB 901 AAACCTCCAAAGAGACTGTGACCTCAAGCCAGAAACAAAGTCTTTCACTCGAGTGGTT 960
  
```

```

QY 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
DB 961 GATGATGCTTTCTTCTTCCATTAAGGCTCTTAATGATCTCATGTGTCTCAAGAAAGCATTTAAAGA 1020
QY 341 IleProSerIleIleGlyValAsnAsnHISGluCysGlyPheLeuLeuProMetLysGlu 360
DB 1021 ATTCTTCATCATCATGAGTGAATCAATTAACAGAGTGTGCTTCTCTGCTCATTAAGAGAG 1080
QY 361 AlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHISLeuIleGlnAsnIle 380
DB 1081 GCTCCTGAGATCTTCAAGTGGCTCCCAAGAGTCCCTTGGCCCTTCATGATTAACAAACATTC 1140
QY 381 LeuHISLIEProProGlnTrpLeuHISLeuValAlaAsnGluTrpPheHISAspLysHIS 400
DB 1141 CTGCACATCCCGCTCAATTTTGCACTTTGTGCTTAATGAATCTTCATCAACAGACAC 1200
QY 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
DB 1201 TCCCTGACTGAATCCGAGACAGTCTTGGACTTGGAGATGTGTTCTTTGTGGTC 1260
QY 421 ProAlaLeuIleThrAlaArgTrpHISArgAspAlaGlyAlaProValTrpPheTrpGlu 440
DB 1261 CCTGCACATCAACAGCTCGATATCAAGAGATGCTGGTGCACCTGTCTACTTATAGAG 1320
QY 441 PheArgHISArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHIS 460
DB 1321 TTTCCGACCCGGCTCAAGTCTTGAAGACAGAGCCGGCTTTGTCAAGCCAGCACAC 1380
QY 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
DB 1381 GCTGATGAAGTCCGCTTTGTGTGGTGTGCTTCTGAAGAGGGGACATTTGTATGTTTC 1440
QY 481 GluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTrpAlaThr 500
DB 1441 GAAGGAGCCACGAGAGAGAGAAATTAAGTCAAGCCGGAAGATGAATAAATACGGGCTACC 1500
QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
DB 1501 TTGCTCCAAACCGGAAATCTTAATGGAAACAGCTGTCTGTGGCCAGCTTAATATCTG 1560
QY 521 ThrGluGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
DB 1561 ACTGAGAGTACCTTCACCTGAGCTTAATGAACATGAACCTCGAGACAGACTCAAGAACCG 1620
QY 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHIS 560
DB 1621 CGGGTGAATTTTGGACAGACCAATCCCTGATCTGTGCTGCCGACATGCTCCAC 1680
QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
DB 1681 AGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY 581 Pro 581
DB 1741 CCT 1743
  
```

RESULT 3
 AR438764 2158 bp DNA linear PAT 20-FEB-2004
 LOCUS AR438764
 DEFINITION Sequence 1 from patent US 6664091.
 ACCESSION AR438764
 VERSION AR438764.1 GI:42663739
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES 1 (bases 1 to 2158)
 AUTHORS Curtis,R.A.J. and Silos-Santiago, I.
 TITLES 53010, a human carboxylesterase family member and uses thereof
 JOURNAL Patent: US 6664091-A 1 16-DEC-2003;
 FEATURES Location/Qualifiers
 source 1..2158

ORIGIN /organism="unknown"
/mol_type="genomic DNA"

Alignment Scores:
Pred. No.: 4,08e-252 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-023-515-2 (1-581) x AR438764 (1-2158)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
DB ATCCACAGGAGACTTACTTATCTGCTTCACAAATGGCTTTTCCGTGATTCACGCC 155

QY 21 LeuLeuGlyYHIsaArgInTrpGlyLysThrGlyProSerAlaGluGlyProGlnArgAsn 40
DB CTGTGGGACACAGACAGTGGGGAATACTGGGCTTCCTGTCAGAGGCCACAGAGAAC 215

QY 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
DB ACCAGGCTGGAGATGATTCAGAGGCAAGTCACCTGTCTGGGAAGCCCTGTGCTGTG 275

QY 61 AsnValPheLeuGlyValProPheAlaProProLeuGlySerLeuArgPheThrAsn 80
DB AACGTGTTCTCTGGAGATCCCTTTGCTGCTCCCGCTGGATCCCTGGATTTACGAA 335

QY 81 ProGlnProAlaSerProTrpPhePheLeuArgGluAlaThrSerTyrProAsnLeuGly 100
DB CCGAGCCCTGCATCGCCCTGGATTACTTCGAGAAAGCACCTCTACCTTAATTTGTC 395

QY 101 LeuGlnAsnSerGlyTrpLeuLeuLeuAspGlnHIsMetLeuLysValHIsTyrProLys 120
DB CTCACAGACTCAGAGTGGCTCTCTTAAGTCAACACATGCTCAAGTGCATTTACGAA 455

QY 121 PheGlyValSerGlyAspCysLeuTyrLeuAsnIleTyrAlaProAlaHIsAlaAspThr 140
DB TTCGAGAGTCTCAGAAAGACTGCTTACCTGAACATCTATGGCCCTCCAGCCGATACA 515

QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
DB GGCTCAAGCTCCCGCTGTGGTGGTTCACAGAGGCTGCTTCACAACTGCTGACCC 575

QY 161 SerIlePheAspGlySerAlaLeuAlaTyrGlyAspValLeuValValValGln 180
DB TCATCTTTGATGGGTCCGCCCTGGCTGCTCATGAGACGTGGTGTGTGTGCTCAG 635

QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHIsAlaProGlyAsnTrp 200
DB TACGGGCTAGGAATAATTTGGTTTCTTCACCATGGGATCAGCATGCTCCGGGAACTGG 695

QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGluPhePheGly 220
DB GCCTTCAAGACACAGTGGCTGCTGTCTGGGTCCAGAAACATCGAGTCTTGCTG 755

QY 221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer 240
DB GGGGACCCCGAGCTGTGACCATTTGGCGAGTCCCGGGAGCATTAATGTTTCTAGT 815

QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHIsLysAlaIleMetGlySerGlyVal 260
DB CTTTATATCTGCTCCCATGGCCAAAGGCTTAATTCACAAACCATCAGAGAGTGGGGTG 875

QY 261 AlaIleIleProTyrLeuGluAlaHIsAspTyrGlyLysSerGlyAspLeuGlnVal 280
DB GCCATCATCCCTTAACCTGAGGCCCATGATTATGAGAGTGAAGCTTCGACGGTGGT 935

QY 281 AlaHisPheCysGlyLysAsnAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThr 300
DB GCACATTTCTGTGTAAACAATGGCTCAGACTGAGGCCCTGTGAGGTGCTCAGAGACA 995

QY 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVal 320
DB AAACCTCCAAAGAGCTGTGACCTTCAGCCAGCCAGAAACAAAGTCTTTACATCGAGTGT 1055

QY 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysValaPheLysAla 340
DB GATGGTGTCTTTCTTCTTAATGAGCTCTAGATGATTTGTCAGAAAGCATTTAAACA 1115

QY 341 IleProSerIleIleGlyValAsnAsnHIsGluCysGlyPheLeuLeuProMetLysGlu 360
DB ATTCCTTCATCATGTGAGTAATTAACACAGAGTGTGCTTCCGTGCTCCATGACAGAG 1175

QY 361 AlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHIsLeuIleGlnAsnIle 380
DB GCTCCTGAGATCCTCAAGTGGCTCCAAAGTCCCTTGCTTCATCTGATACAAACATC 1235

QY 381 LeuHIsIleProProGlnTrpLeuHIsLeuValAlaAsnGluTyrPheHIsAspLysHIs 400
DB CTGACATCCCGCTCAGTATTTGACCTTGTGGCTAATGAAATCTTCATGACAAACAC 1295

QY 401 SerLeuThrGluIleAspAspSerLeuLeuAspLeuGlyAspValaPhePheValVal 420
DB TCCCTGACTAAATCCAGACAGTCTTCTGACCTTGTGAAGATGTGTTCTTTGTGTC 1355

QY 421 ProAlaLeuIleThrAlaArgTyrHIsArgAspAlaGlyAlaProValTyrPheTyrGlu 440
DB CCTCCACTGATCACAGCTCAGATATCACAGAGATGTGTGTCACCTGTCTATGAG 1415

QY 441 PheArgHIsArgProGlnCysPheGlyAspThrLysProAlaPheValLysAlaAspHIs 460
DB TTTTGGACCGGCTCAGTCTTGAAGACACAAACCGGCTTTTGTCAAGCCGACAC 1475

QY 461 AlaAspGluValaArgPheValaPheGlyGlyAlaPheLeuLysGlyAspIleValaMetPhe 480
DB GCTGATGAATCCGCTTTGTGTTCGGTGGTCTTCTGTAAGGGGACATGTTAATGTC 1535

QY 481 GluGlyValaPheThrGluGluGlyLysLeuSerArgLysMetLysTyrTrpAlaThr 500
DB GAAAGAGCCACGGAGAGGAGAAATTACTGAGCCGGAATGATGAATACTGGCTACC 1595

QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeu 520
DB TTTGCTCGAAACCGGAAATCTTAATGGAAAGACCTGTCTGTGGCAGACTTAATATCG 1655

QY 521 ThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
DB ACTGACAGTACTCCAGCTGGAATTGAACATGAGCTCGACAGAGACTCAAAAGACCG 1715

QY 541 ArgValAspPheThrPheSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHIs 560
DB CGGGTGAATTTTGGACACAGACACATCCCTGAATCTGTCTGCTCGACATGCTCAC 1775

QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
DB AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835

QY 581 Pro 581
DB 1836 CCT 1838

RESULT 4
AX600124 2158 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 1 from Patent WO0250256.
DEFINITION
ACCESSION AX600124
VERSION AX600124.1 GI:28400190
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1

AUTHORS Curtis, R.A. and Silos-Santiago, I.
TITLE 53010, a novel human carboxylesterase family member and uses thereof
JOURNAL Patent: WO 0250256-A 1 27-JUN-2002;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES
 source
 1. 2158
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 96. .1841
 /note="unnamed protein product"

CDS

/codon_start=1
 /protein_id="CAD66643.1"
 /db_xref="GI:28400191"
 /translation="MPQGLTSSASQWCFLLIQPLIGHKQKCKPSABGPRNTRLG
 MIQKQVTLSPVFNVPFLVFPAPAPLGSRLFTNPSPASWDLREKSTPNICLQ
 NSEMLDQHLKHYPRKGVSEDLVYNIYAPAHADPSKLPVLVWPPGAPKFGSA
 SIPDSALAAVEDLVVVVQYRLGIFGETTWDQHPGNMVAFKDQVALSWQKNEF
 FGDPSSVTIRFESAGAISVSLSIPAKGLFKAIMESGVAILPYLEADHYKESD
 LQVVAHFGNNADEALRLCRTPSKELLTLQKTSFTRVNDGAPPNPDLILS
 QKAPRAISIIIGVNNHCGFLIPKKAPEIISGSKSLALHLIQNTLHIFPOLYHLVA
 NEYFDKSLTEIRDSLDLDLGDVFPVVPALITARYHDDAGAPVYFERRHPQCFED
 TKPAVVKADHADVAFVFGAFLKGDIVMFEGATEERKLSIKMKKYATFARTGNPN
 GNDLSIMPAYNLTEQYLQDLNMSLGRLKEPRVDFMTSTIPLILASDMLHSPLSL
 TFLSLQRPFFCAP"

ORIGIN**Alignment Scores:**

Pred. No.: 4,086-252 Length: 2158
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-10-023-515-2 (1-581) x AX600124 (1-2158)

QY 1 MetProGInGlyLeuThrSerSerAlaSerGInTPCyRPhPheLeuLleLeuGInPro 20
 Db 96 ATGCCACAGGAGCTTACTTCATCTCTTCAACATGATGCTTTTCTGATTTCTCAAGCCC 155
 QY 21 LeuLeuGInLysArgGInTPGlyLysThrGlyProSerAlaGInGlyProGInArgAsn 40
 Db 156 CTGTTGGGACACAGCAAGTGGGGAAACTGGGCTTCTCGCTGAAGGCCACAGAGAAC 215
 QY 41 ThrArgLeuGlyTTPLeuGInGlyLysGInValThrValLeuGlySerProValProVal 60
 Db 216 ACCAAGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTG 275
 QY 61 AsnValPheLeuGlyValProPheAlaAlaProPheLeuGlySerLeuArgPheThrAsn 80
 Db 276 AACGTGTCTCTGGAGTCCCTCTTGTCTGCCCGCTGGGATCCCTGGGATTTACGAAAC 335
 QY 81 ProGInProAlaSerProTrpAspAsnLeuArgGInLysAlaThrSerTyrProAsnLeuGly 100
 Db 336 CCGCAGCTGTGATCGCCTTGGGATTACTTGCAGAAAGCACTCTTCACTTAATTTGTGC 395
 QY 101 LeuGInAsnSerGInTTPLeuLeuLeuAspGInHisMetLeuLysValHisTyrProLys 120
 Db 396 CTCAGAACTGAGAGTGGCTGCTTTAAGTCAACACATGCTCAAGTGTGATTAACCGAAA 455
 QY 121 PheGlyValSerGInAspCysLeuTyrLeuAsnLleTyrAlaProAlaHisAlaAspThr 140
 Db 456 TTCGAGAGTGTGAGAACTGCTCTTACCTGAACATCTATGCGCTGCCCAACCCCATCA 515
 QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
 Db 516 GAGCTCAAGCTCCCGCTTGTGTGTGTCCAGAGAGGCTCTTCAAGAGCTGGCTAGCG 575
 QY 161 SerLysPheAspGlySerAlaLeuAlaAlaTyrGlyAspValLeuValValValGln 180
 Db 576 TCCATCTTTGATGGGTCCGCGCTGGCTGTATGAGAGAGTGTCTGTGTGTGTGTGTG 635

QY 181 TyrArgLeuGlyLysPheGlyPhePheThrThrTrpAspGInHisAlaProGlyAsnTyr 200
 Db 636 TACCGGCTAGGAATATTTGGTTTCTTCAACACATGGGATCAGACGTCCGGGAACTGC 695
 QY 201 AlaPheLysAspGInValAlaAlaLeuSerTrpValGlnLysAsnLleGlnPheGly 220
 Db 696 GCTTCAAGAGACAGAGTGGCTGTCTCTCGGGCTCCAGAAAGACATGAGTTCCTCGGT 755
 QY 221 GlyAspProSerSerValThrLlePheGlyGlySerAlaGlyAlaLleSerValSerSer 240
 Db 756 GGGAGCCCAAGCTGTGATCAATCTTTGGCAGTCCGGGAGCCATAGGTTCCTAGT 815
 QY 241 LeuLleLeuSerProMetAlaLysGlyLeuPheHisLysAlaLleMetGlySerGlyVal 260
 Db 816 CTTATACTGTCTTCCATAGGCCCAAGGCTTATTCACAAAGCATATGAGAGTGGGGT 875
 QY 261 AlaLleLleProTyrLeuGInLysAlaHisAspTyrGlyLysSerGlyAspLeuGInVal 280
 Db 876 GCCATCATCCCTTACCTGAGGCCCATGATTATGAGAAAGTGAAGACCTGCAAGGTGTT 935
 QY 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThr 300
 Db 936 GCACATTTCTGTGTACCAATGCGTCAGACTCTGAGGCCCTGTGAGGTGCTGAGACA 995
 QY 301 LysPProSerLysGlyLeuLeuThrLeuSerGInLysThrLysSerPheThrArgValVal 320
 Db 996 AACCTCTCAAGAGACTGTGACCTTCAAGCCAGAAACAAAGTCTTTCCTCTGAGTGGTT 1055
 QY 321 AspGlyAlaPhePheProAsnGInProLeuAspLeuLeuSerGInLysAlaPheLysAla 340
 Db 1056 GATGTGTCTTCTTCTTCCATAGAGCTCTTAATGATCTATTGTCTCAAGAAACATTTAAGA 1115
 QY 341 IleProSerLleLleGlyValAsnAsnHisGlyCysGlyPheLeuLeuProMetLysGly 360
 Db 1116 ATTCCTTCATCATCAGGATCAATTAACACAGAGTGTGCTTCTGCTCTCATATGAAGAG 1175
 QY 361 AlaProGInLleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuLleGlnAsnLle 380
 Db 1176 GCTCTGAGATCTCTAGTGGCTCAACAAAGTCTTGTGCTTCATCTATACAAACATTC 1235
 QY 381 LeuHisLleProProGInTyrLleuHisLeuValAlaAsnGInTyrPheHisAspLysHis 400
 Db 1236 CTGACATCCCGGCTCAATTTTGGACCTTGAGGTATTAATCTTCATGACAAAGAC 1295
 QY 401 SerLeuThrGInLleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
 Db 1296 TCCCTGACTGAATCCGAGACAGTCTTCTGACCTTGTGAGAGTGTCTTGTGTGTC 1355
 QY 421 ProAlaLeuLleThrAlaArgTyrHisAspAspAlaGlyAlaProValTyrPheTyrGln 440
 Db 1356 CCTGACTGATCAACAGCTCGATATCAAGAGATGCTGTGACCTGTACTCTTATAGAG 1415
 QY 441 PheArgHisAspProGInCysPheGlyAspThrLysProAlaPheValLysAlaAspHis 460
 Db 1416 TTTCGGCACCGGCTCAATGCTTTGAAGACAGAAAGCGGCTTTGTCAAGCCGACAC 1475
 QY 461 AlaAspGInValArgPheValPheGlyGlyAlaPheLeuLysGlyAspLleValMetPhe 480
 Db 1476 GCTGATGAAGTCCGCTTGTGTGTGTGCTGCTTCTGAAAGGGGAGCATTTGTATGTTTC 1535
 QY 481 GlnGlyValaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTyrTrpAlaThr 500
 Db 1536 GAAGGAGCCAGAGAGAGAGAGATTAAGTGAAGCGGAAAGATGAATAATCTGGCTACCC 1595
 QY 501 PheAlaArgThrGlyAspProAsnGlyAsnAspLeuSerLeuTTPProAlaTyrAsnLeu 520
 Db 1596 TTTCGTCAACCGGGAATCTTAATGGAAACGACTGTCTGTGGCCAGCTTAATATG 1655
 QY 521 ThrGlnGInTyrLeuGInLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
 Db 1656 ACTGAGAGTACCTTCAGCTGAGCTTGAACATGAGCCTCGGACAGAGACTCAAGAAACCG 1715

QY 541 ArgValAspPheThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
Db 1716 CGGGTGGATTTTGGACACACCATCCCTGATCTGTCTGCTCCACATGCTCCAC 1775
QY 561 SerProLeuSerSerIleuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
Db 1776 AGTCCCTCTTCTTCTTCACTTACTTCTCTCTCTCTCCACGCTTCTTCTTCTTGTGCT 1835
QY 581 Pro 581
Db 1836 CCT 1838
RESULT 5
LOCUS BC069501
DEFINITION Homo sapiens hypothetical protein FLJ1547, mRNA (cdna clone
ACCESSION BC069501 2047 bp mRNA linear pri 19-AUG-2004
VERSION BC069501 MGC:97182 IMAGE:7262427), complete cds.
KEYWORDS MGC.
SOURCE BC069501.1 GI:46854682
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 2047)
Strusberg,R.D., Felngold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,C.M., Shennen,C.M., Schuler,G.D.,
Altchul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schert,T.B., Brownstein,M.J., Uedlin,T.B., Toshiyuki,S.,
Carinini,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., Mcwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huliy,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Pahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywanski,M.I., Skalska,U., Smalls,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2047)
DIRECTOR MGC Project.
TITLE Submitted (29-APR-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Baylor Human Genome Sequencing Center
cDNA Library Preparation: Baylor Human Genome Sequencing Center
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Louised, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIN at: http://image.llnl.gov
Series: IRBR Plate: 4 Row: e Column: 1.
Location/Qualifiers

source
1. 2047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:97182 IMAGE:7262427"
/clone_type="PCR rescued clones"
/clone_id="NH_MGC_277"
/note="Vector: pPCR-Script Amp SK(+) with reversed insert"
1. 2047
/gene="FLJ1547"
/db_xref="locusID:221223"
121. 1848
/gene="FLJ1547"
/codon_start=1
/product="FLJ1547 protein"
/protein_id="AAH69501.1"
/db_xref="GI:46854683"
/db_xref="locusID:221223"
/translation="MSGNMYHPQIILMAIYVLAAPKPSAEGPQANTLMIQKQ
VTUIGSPVNVNPELVFPAPPLGSLFTTPQASPDNDLREATSYENLCLQSEWIL
LDQMLKVHPTKGVSESDCLYLNIYAPAHADTGSKLPLVWFGAFKXGASIPDS
ALAAVEDLVVVVQYRLGIFFTWDQAPGMWAFQDVVALSWQKNEPFGDBS
SVTIFGSAGALISVSLILSPMAKGLPHKAIMSGVALIPLYLAHYEKSEDLQVVAH
FCGNMADSSEALRLCRLTKTSKELTSLQKTSKTRVVDGAFPPNPLDLSOKAFKA
IPSLIGVNNHECGFLPMKEAPRILSGNSKLALHILONTIHTPQYHILVANEYFD
KSHLRDRSLDLGLDVPFVVPALITAVYHRNAGAPVYTFERRHPOCFEDTKPAFV
KADHADRVRVFPAGAPLKDIIVFEGATEBKLSKMMKYATFARTGPNNDLSL
WPNVNLTEQLDLDLNMISLQQRKEPRVDFWSTIPLILASDMLHSPSLTFLSL
QPFEPFCAP"
ORIGIN
Alignment Scores:
Pred. No.: 3 486-238 Length: 2047
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservative: 3
Best Local Similarity: 97.02% Mismatches: 6
Query Match: 94.67% Indels: 8
DB: 9 Gaps: 1
US-10-023-515-2 (1-581) x BC069501 (1-2047)
QY 12 TrpCysPhePheLeuIleLeuGlnProLeuLeuGlnYHlaArgGlnTrpGly 31
Db 160 TGGGCTATCTGGGCTCTTGGACGCC-----ACCAAGG 195
QY 32 ProSerAlaGlnGlyProGlnArgaenThrArgLeuGlyTrrpIleGlnGlySerGlnVal 51
Db 196 CTTTCTGAGAGGGCCACAGAGAACACACAGCTGGATGATTCAGGACGACAGTGC 255
QY 52 ThrValLeuGlySerProValProValaenValPheLeuGlyValProPheAlaPro 71
Db 256 ACTGTCTGGAGAGCCCTGTGCTGACGCTGTCCTGAGATCCCTTCTGCTGCC 315
QY 72 ProLeuGlySerLeuArgPheThrAnPProGlnProAlaSerProTrpAspAnLeuArg 91
Db 316 CGGCTGGATCCCTGGATTTACGAACCGCAGCTGCATCCCTCGGAGTAACTTGCGA 375
QY 92 GluAlaThrSerTrpProkneuCysLeuGlnaenSerGluTrpLeuLeuLeuSerGln 111
Db 376 GAAGCACCCTTACCTTAATTTGGCTTCAGAACTCAGAGTGGCTGCTCTTAATCA 435
QY 112 HisMetLeuValaHisTrpProLysPheGlyValaSerGluAspCysLeuTrpLeuAn 131
Db 436 CACATGCTCAAGTGCATTAACCGAATTCGAGTGCAGAACTGCTGCTACTGAAAC 495
QY 132 IleTrpAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhePro 151
Db 496 ATCTAGCGCTGCCACGCGATACAGCTCCAGCTCCCTCTTGTGTGGTTCCTCA 555
QY 152 GlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaTrp 171
Db 556 GGAAGTGCCTTCAAGACTGCTCAGCTCCATCTTTATGGGTCCGCTGCTGCTAT 615

Qy	172	GLuAerPvAlLeuValValValValGlnTrpTrpArgLeuGlyIlePheGlyPhePheThr	191
Db	616	GAGAGCGGTGGTGTGTGTGTGTCCAGTACCGGCTAGGAATAATTTGGTTCTTACCA	675
Qy	192	TrpAerPGLnHIsaIAProGlyAAsnTrpAlaPheLysAerGlnValAlaAlaLeuSerTrp	211
Db	676	TGGATCATGCATGCTCCCGGGAGACCTGGGCTTCAAGACCAAGTGGCTGTCTGTCCGG	735
Qy	212	ValGlnLysAAsnIleGluPhePheGlyGlyAAspProSerSerValThrIlePheGlyGlu	231
Db	736	GTCCAGAAAGACATCATAGATTCTTCGGTGGGAGCCCAAGCTGTGATCCATCTTTGGCAG	795
Qy	232	SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPhe	251
Db	796	TCCCGGGAGCGCAATAGTGTTCATAGCTTATACGTCTCCCATGGCCAAAGGCTTATTC	855
Qy	252	HisLysAlaAlaIleMetGluSerGlyValAlaAlaIleIleProTrpLeuGlnAlaHIsAAspTrp	271
Db	856	CACAAAGCCATCATAGAGAGTGGGGGTGCCATCATCCCTTACCTGGAGGCCCATGATAT	915
Qy	272	GluLysSerGlnAAspLeuGlnValValAlaHIsAAspCySGlyAAsnAlaIleSerAAspSer	291
Db	916	GAGAAAGATGAGACCTGTCAGTGGTTCATATTTCTGTGTAACAATGCGTCAAGACTCT	975
Qy	292	GluAlaLeuLeuAAspCySGlyLeuAArgThrLysProSerLysGluLeuLeuThrLeuSerGln	311
Db	976	GAGGCCCTGCTGAGAGGTGCTTGAAGCAAAACCTTCCAAAGAGCTGTGACCTCAGCCAG	1035
Qy	312	LysThrLysSerPheThrArgValValAAspGlyAlaPhePheProAAsnGluProLeuAAsp	331
Db	1036	AAACCAAGTCTTTCATCTCGAGTGTATTAGTGTCTTTCTTCTTAATGAGCCTTAGAT	1095
Qy	332	LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAAsnAAsnIleGlu	351
Db	1096	CTATTGTCTCGAAAGACATTTAAAGCAATCTCTTCCATCATCGAGTCAATACCAAGAC	1155
Qy	352	CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAAsnLysSer	371
Db	1156	TGTGGCTTCCGTGCTCCATATAGAGAGGCTCCTGAGATCCTCAGTGGCTCCAAAGTCC	1215
Qy	372	LeuAlaLeuHIsaIleGlnAAsnIleLeuHIsaIleProProGlnTrpLeuHIsaIleVal	391
Db	1216	CTTGGCCCTCCATCGATACAAACCAATCTTGCACATCCGCGCTCAAGATATTGACCTTGTG	1275
Qy	392	AlaAAsnGluTrpPheHIsAAspLysHIsaSerLeuThrGluIleAAspSerLeuLeuAAsp	411
Db	1276	GCTAATGAATATCTTCCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTTGTGAC	1335
Qy	412	LeuLeuGluAAspValPhePheValValProAlaLeuIleThrAlaArgTrpHIsaArgAAsp	431
Db	1336	TTGCTTGAGATGATGTTCTTGTGTGGTCCCTGCATGATCAAGCTCGATATCACAGAGAT	1395
Qy	432	AlaGlyAlaProValTrpPheTrpGluPheArgHIsaArgProGlnCysPheGluAAspThr	451
Db	1396	GCTGTGTGACCTGCTACTTCTTATGAGTTTGGCACCGGCTCATGTGTGTTGAAGACACG	1455
Qy	452	LysProAlaPheValAlaLysPheHIsaAspHIsaIleAAspGluValAAspPheValPheGlyAla	471
Db	1456	AAAGCCGCTTTTGTTCAAAGCCGACCAAGCTAGTAAGTCCCGCTTGTGTTCGGTGTGCC	1515
Qy	472	PheLeuLysGlyAAspIleValIleMetPheGlnGlyAlaThrGlnGlnGluLysLeuLeuSer	491
Db	1516	TTTCTCGAAGGGGGACATGTTATATGTTCCAAAGAGACCGAGAGAGAGAAATTTCTAGAC	1575
Qy	492	ArgLysMetLysLysTrpAlaThrPheAlaArgThrGlyAAsnProAAsnGlyAAsnAAsp	511
Db	1576	CGAAGATGATGAATATACGTGGCTACCTTTGCTCAACCGGGAATCCCTAAATGGAAACAC	1635
Qy	512	LeuSerLeuTrpProAlaTrpAAsnLeuThrGlnGluTrpLeuGlnLeuAAspLeuAAsnMet	531
Db	1636	CTGTCTCTGTGGCCAGCTTATATGTGACTGAGCAGTACCTTCCAGCTGGACTTGAAACATG	1695
Qy	532	SerLeuGlnGlnAArgLeuLysGluLeuProAArgValAAspPheTrpThrSerThrIleProLeu	551

Dd	1696	AGCTTCGGACGACACTCAAGAACC	CGGTGATTTTGGACCAACCATCCCTG	1755
Oy	552	IleuSerAlaSerAapMetLeuHisSerProIeuSerIeuThrPheIeuSerLeu		571
Dd	1756	ATCCTGTCGTGCCTCGACATGCTCCACAGTCCTCTTCTTCTTAACCTTCTCTCTC		1815
Oy	572	LeuGlnProPhePhePhePheCyValAlaPro	581	
Dd	1816	CTCCAGCCTTCTTCTTCTTCTTGCTCT	1845	
RESULT 6				
LOCUS	AX477711			
DEFINITION	Sequence 23 from Patent WO0246426.			
ACCESSION	AX477711			
VERSION	AX477711.1			
KEYWORDS	GI:22216866			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			
AUTHORS	Sanjawaia,M.M., Yao,M.G., Au-Young,J., Baughn,M.R., Arvizu,C.,			
	Ring,H.Z., Lee,E.A., Ding,L., Hafalala,A.J., Tang,Y.T., Yue,H.,			
	Tribouley,C.M., Lu,D.A., Lal,P.G., Warren,B.A., Yang,J.,			
	Walia,N.K., Nguyen,D.B., Gandhi,A.R. and Ison,C.H.			
TITLE	Drug metabolizing enzymes			
JOURNAL	Patent: WO 0246426-A 23 JUN-2002;			
	Incyte Genomics, Inc. (US)			
FEATURES	Location/Qualifiers			
Source	1..2232			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	/note="Incyte ID No: 653808CB1"			
ORIGIN				
Alignment Scores:				
Pred. NO.:	7,39e-237	Length:	2232	
Score:	2900.00	Matches:	548	
Percent Similarity:	99.82%	Conservative:	2	
Best Local Similarity:	99.46%	Mismatches:	1	
Query Match:	94.19%	Indels:	0	
DB:	6	Gaps:	0	
US-10-023-515-2 (1-581) x AX477711 (1-2232)				
Oy	31	GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValGln	50	
Dd	274	GGGCCTTCTGCTGAAGGGCCACAGAGGAACAACCGCTGGATGATTCAGGGCCAAACCA	333	
Oy	51	ValThrValIleuGlySerProValProValAsnValPheIleuGlyValProPheAla	70	
Dd	334	GTCACGTGTGGGAAACCCCTGTGCTGTGAACGTGTTCTCGGAGTCCCTTGTGCTCT	393	
Oy	71	ProProIeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTirPaspAsnLeu	90	
Dd	394	CCCCCGCTGGGATCCCGCATTTTAGAACACCCGAGCCTGCATGCCCTGGGATACTTG	453	
Oy	91	ArgGluAlaThrSerTyProAsnLeuCylLeuGlnAsnSerGluTTPleuLeuLeuAsp	110	
Dd	454	CGAAGAACCACTCTCTACCTTAATTGTGCTCCAGAACTCAGAGTGCTGCTTTAGAT	513	
Oy	111	GlnHisMetLeuLysValHisTyProValPheGlyValSerGluAspCybleuTyTLeu	130	
Dd	514	CACACACATGCTCAAGGAGCATTTACCGAAATTCGAGATGTCAGAAAGCTGCTTACTG	573	
Oy	131	AsnIleTyAlaAlaProAlaHisAlaAspThrGlySerTyblauProValIleuValTirPhe	150	
Dd	574	AACATCTATGCGCTGCGCACGCGGATACAGGGCTCCAAGCTCCCGTCTTGTTGGTTC	633	
Oy	151	ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAla	170	

[illegible]

Db	1714	GACCTGTTTCTGTGGCCAGCTTATATCATGACTGACAGTACCTCAGCTGACTGTAAC	1772
Qy	531	MeSeSerLeuG1g1nArgLeuYsg1uPProArxVa1aapPhetRThSerThrllePro	550
Db	1774	ATGAGGCTCGGACAGACTCAAAAGAACCGGGGTGGAGTTTGGACGACGACCATCC	1833
Qy	551	Leu1leLeuSerAlSerAspMetLeuHisSerProLeuSerSerLeuThPhleuSer	570
Db	1834	CTGATCTCTGTGCTCGCTCCGACATGCTCCACAGTCTCTTCTTCTTAATTCTCTCT	1893
Qy	571	LeuLeuG1nProPhaPhePhePheCysAlaPro	581
Db	1894	CTCCTCAGCCTTTCTTTTCTTTTGTGCTCT	1926
RESULT 7		BC039073	2298 bp mRNA linear PRI 18-NOV-2002
LOCUS		BC039073	Human sapiens, similar to carboxylesterase, clone IMAGE:5174937,
DEFINITION		BC039073	mRNA.
ACCESSION		BC039073	
VERSION		BC039073.1	GI:25056132
KEYWORDS			
SOURCE		Human sapiens (human)	
ORGANISM		Human sapiens	
REFERENCE		Mumukya, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mumukya, Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
TITLE		1 (bases 1 to 2298)	
JOURNAL		Strausberg, R.	
REMARK		Submitted (01-NOV-2002) National Institutes of Health, Mammalian	
COMMENT		Gene Collection (MGC), Cancer Genomics Office, National Cancer	
		Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,	
		USA	
		NIH-MGC Project URL: http://mgc.nci.nih.gov	
		Contact: MGC help desk	
		Email: cgapbs-remail.nih.gov	
		Tissue Procurement: Life Technologies, Inc.	
		cDNA Library Preparation: Life Technologies, Inc.	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
		DNA Sequencing by: Baylor College of Medicine Human Genome	
		Sequencing Center	
		Center code: BCM-HGSC	
		Web site: http://www.hgsc.bcm.tmc.edu/cdna/	
		Contact: amg@bcm.tmc.edu	
		Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsesed, H.,	
		Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,	
		A.N., Gibbs, R.A.	
		Clone distribution: MGC clone distribution information can be found	
		through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov	
		Series: IRAC Plate: 51 Row: k Column: 18	
		This clone was selected for full length sequencing because it	
		passed the following selection criteria: Hexamer frequency ORF	
		analysis.	
FEATURES		Location/Qualifiers	
SOURCE		1..2298	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5174937"	
		/tissue_type="Brain, Lung, Testis, adult, pooled whole"	
		/clone_lib="NIH MGC_115"	
		/lab_host="DH10B"	
		/note="Vector: pCMV-SPORT6"	
ORIGIN			
Alignment Scores:			
Pred. No.:	2, 89e-235	Length:	2298
Score:	2881.50	Matches:	555
Percent Similarity:	90.85%	Conservative:	1
Best Local Similarity:	90.69%	Mismatches:	7
Query Match:	93.59%	Indels:	49
	9	Gaps:	3

US-10-023-515-2 (1-581) x BC039073 (1-2228)

QY 14 RHEPHELEU1LEUENINPROLEUENUGLYHISARGINTPTGILYUERTHRTGILYPROSER 33
 DB 67 TTTTTCCTGCTCTTTGGATAAA-----AACCACTTTAACTGG-----GGGCTTTCT 111
 QY 34 AIAIGLUGLYPROGILNARGANETHARGLEUGLYTPTILGLNGLYLVSGLNVALTHRVAL 53
 DB 112 GGTGAAGGGCCACAGAGAACACAGCTGGATGATTCAGGGCAAGCAAGTCACTG 171
 QY 54 LEUGLYSERPROVALPROVALANVALPHELEUGLYVALPROPHEALAAAPROPROLEU 73
 DB 172 CTGGGAAGCCCTGCTGCTGGAACGTGTTCTCGAGAGTCCCTTTGCTCTCCCGCTG 231
 QY 74 GLYSERLEUARGPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 93
 DB 232 GGATCCCTGCGATTAAGAACCCGAGCTGATCCCTGGATTACTTGGAGAGAGCC 291
 QY 94 THRSERYPROASMLEUCYALEUGLNAANSERGLUTPTLEULEUENAPGILNHSMET 113
 DB 232 ACCTCTACCTTAATTTTGCTCTCCAGAACTCAGAGTGGCTGCTTAATCAACAGT 351
 QY 114 LEULYSVALHISERYPROLYSPHEGLYVALSERGLUASPCELEUTYRLEUENILERY 133
 DB 352 CTCAAGGTGCATTACCCGAAATTCGAGGTGTCAGAAAGATGCTGCTACCTGAACATCTAT 411
 QY 134 AIAAPROALHISALASPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 153
 DB 412 GGGCTGCGCCACGCGATTCAGGCTCCAGGCTCCCGCTGTTGGTGGTTCACAGAGAGGT 471
 QY 154 AIAPELYSTHRTGLYSERLASERILEPHEAPGILYSERIALALEUALATYRGLUASP 173
 DB 472 GCTTCAGAGCTGGCTCAGCTCTCCATCTTGATGGGTCCGCTGCTGCTCATAGAGAC 531
 QY 174 VALLEUVALVALVALVALGINTYRARGLEUGLYLEPHEGLYPHEPHEPHEPHEPHEPHE 190
 DB 532 GTCGTGTTGGTGTGTCAGTACCGGCTAAGAAATTTGGTTCTTCACTGAGTCTCT 591
 QY 190 ----- 190
 DB 592 CCMAATGATGACATCCAGATCCCAAGAGGAGAGTGCATGACCATAGACAGTGC 651
 QY 190 ----- 190
 DB 652 CAGCAACTTAAGAGAGGAGGAGCTTCGGGACATGAACAAACCTGGCCCTTAATCTT 711
 QY 191 ---THRTTPASPGILNHSALAPROGLYASNTPTALAPHELYASPGLNVALAIALEU 209
 DB 712 AGCAGATGGGATCAGATGCTCCGGGAACTGGGCTTCAGAGACAGGTGGCTGCTG 771
 QY 210 SERTTPVALGLNLYAENILEGLUPHEPHEGLYGLYASPPROSERSERVALTHRIEPHE 229
 DB 772 TCCTGGGTCCAGAGAACATCAGATCTTCGGTGGGGAACCCAGAGCTGTCACATCTTT 831
 QY 230 GLYGLUSERALAGLYALALLESERVALSERSERLEULEUSERPROKMERIALYSGLY 249
 DB 832 GCGAGTCCGCGGAGCCATAAGTGTTCATCTTAATCTGCTCCCAATGCGCAAGGCT 891
 QY 250 LEUPHEHISLYALILEMERGLUSERGLYVALAIALEILEPROCTYRLEUGLUALHIS 269
 DB 892 TTATTCACAAAGCCATCATGAGAGTGGGCTGGCCATCATCTTAAGAGGCGCAT 951
 QY 270 AAPTLYRGLYSPSERGLUASPGLNVALVALAIAHISPEHCYGLYANANALASER 289
 DB 952 GATTTATGAGAGAGTGAAGCTGCGAGGTGTTGCACTTCTGTTGTAACATGCTGCTCA 1011
 QY 290 AAPPSEGLUALALEULEUARGCYALEUARGTHTYRSPSERLYSGLEULEUENHRLU 309
 DB 1012 GACTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTTCAGAGAGTGTGACCTTC 1071
 QY 310 SERGLNLYSTHLYSERPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 329

DB 1072 AGCCAGAAAACAAAGCTTTTCACTGAGGTGTGATGCTGCTTTCTTAATAGACCT 1131
 QY 330 LEUAPPLEULEUSERGLNLYSALAPHELYSALALEPROSERILEGLYVALAENAN 349
 DB 1132 CTAGATCTATGTCTCAGAAACATTTAAAGCAATTCCTTCAATCATGGAATCAATAC 1191
 QY 350 HISGLUCYSGLYPHELEULEUPROMETLYSGUALAPROGLUULEUSERGLYSERAN 369
 DB 1192 CACAGGTGTGCTTCTGCTGCTGCTTAAGAGAGGCTCCGATCATCTGATGCTGCACAC 1251
 QY 370 LYSERLEUALALEUHSILEU1LEUENILEUHSILEPROPROGLINTYRLEUHS 389
 DB 1252 AAGTCCCTTGCCCTCATCTGATACAAACATCTGCATCCGCTTCAAGATTTGCAAC 1311
 QY 390 LEUVALAIAENGLTYRPHENISAPLYSNISSERLEUTHGLU1LEARGASERLEU 409
 DB 1312 CTGTGGCTTAATGAATPACTTCCATGACAGACCTCCCTGACTGAANATCCGAGACCTTT 1371
 QY 410 LEUAPPLEULEUGLYASPVALPHEPHEVALPROALALEU1LETHRALAARGTYRHIS 429
 DB 1372 CTGAGCTTGCTGAGAGATGTCTTTGTGTGCTCCCTGACATGATCAGCTGATATCAC 1431
 QY 430 ARGASPVALAGLYALAPROVALTYRPHETRYGLUPHEARGHISARGPROGLNCSYPHEGLU 449
 DB 1432 AGAGATGCTGTGCACTGTCTAATGATGATTTGCGCACCGGCTTCAGGTGCTTGA 1491
 QY 450 AAPTHTYRPROVALPHEVALLYSALASPHISALASERGLNVALARGPHEVALPHEGLY 469
 DB 1492 GACAGAGAGCCGCTTTGTCAAGCCAGCACGCTGATGAGTCCGCTTGTGTGGGT 1551
 QY 470 GLYALAPHELEULYSGLYASPVILLEVALMETPHEGLUGLYALATHRTGLUGLULYSLEU 489
 DB 1552 GGTGCTTCTCAAGGGGACATTTGTATGTTCAAGAGACCAAGAGAGAGAGTA 1611
 QY 490 LEUSERARGLYMETMETIERYTRPALATHRPHENALARGTHRTGLYASNPROMENGLY 509
 DB 1612 CTGAGCCGGAATGATGAATAATGCTGCTAATCTTGTGCTCAACCGGAAATCTAATGG 1671
 QY 510 AENAAPLEUSERLEUTRPROVALTYRASNLEUTHRTGLUINLYRLEUGLNUASPLEU 529
 DB 1672 AACGACCTGTCTGTGCGACCTTAATATGACTGAGCACTACCTGACGCTGAGCTTG 1731
 QY 530 AENMETSERLEUGLYLARGLEULYSGUPROARGVALASPHEPHEPHEPHEPHEPHE 549
 DB 1732 AACATGAGCTTGGACAGAGATCAAGAACCCGGGTGGATTTTGGACGACGACATC 1791
 QY 550 PROLEU1LEUSERLASERASPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 569
 DB 1792 CCCCTGATCTGTGCTGCTCCGACATGCTCCAGATCTCTTCTTCTTAATCTTCCCTC 1851
 QY 570 SERLEULEUGLNPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 581
 DB 1852 TCTCTCTCCAGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1887
 RESULT 8
 AX746724 2092 bp mRNA linear PAT 20-JUN-2003
 LOCUS AX746724
 DEFINITION Sequence 249 from Patent EP1308459.
 ACCESSION AX746724
 KEYWORDS AX746724.1 GI:32130991
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS
 1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
 Yamamoto, J., I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
 Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Negahari, K. and
 Masuno, Y.
 TITLE
 Full-length cDNA sequences
 Patent: EP 1308459-A 249 07-MAY-2003;
 Helix Research Institute (JP) ; Research Association for


```

Db      897  TGCACATTTCTGCTGTAACATGCTGACGCTGAGCCCTGCTGAGGCTGCTGAGAC 956
Qy      300  rlyspProserLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVa 320
Db      957  AAAACCTCCAAAGAGTGTGACCCCTGACCCAGAAAACAAAGCTTTTCACTGAGTGGT 1016
Qy      320  lAeGlyAlaPhePheProAeGluPProLeuAePLeuLeuSerGlnLysAlaPheLysAl 340
Db      1017  TGATGGGCTTTCTTTCTTAATAGCTTCAATCTATTCTCTCAAGAACATTAAAGC 1076
Qy      340  aLlePProserLlelGlyValaLsnaBnHlsgLysGlyPheLeuLeuProMetLysG 360
Db      1077  AATTCCTTCATCATGAGTCAATAGCCACGAGTGGCTTCTCTGCTGCTATGAAAGA 1136
Qy      360  uAlaPProGluLeuLeuSerGlySerSerLysSerLeuAlaLeuLleuLleuLleu 380
Db      1137  GGCTCTTGAGATCTCAAGTGGCTCCAAAGAGTCCCTTGCCCTCATGTATCAAAACAT 1196
Qy      380  eLeuHlslleProProGlnLysLeuHlslleuValAlaAeGluLysPheHlslsl 400
Db      1197  CTTGACATCCCGCTCAGATATTGCACTTGTGGCTAATGAACTTCCATGACAAAGCA 1256
Qy      400  sSerLeuThrGluLleLysAspSerLeuLeuAePLeuLeuGlyAspValaPhePheVala 420
Db      1257  CTCCCTGACTGAATCCGAGACAGCTCTTGACATTTGAGAGATGTGTTCTTGTGGT 1316
Qy      420  lProAlaLeuLleThrAlaArgTyPheHlslslArgAlaGlyAlaPProValLysPheTyG 440
Db      1317  CCTGCACATGATCACAGCTCGATATCACAGAAATGCTGGTGAACCTGCTCACTTATGA 1376
Qy      440  uPheArgHlslArgProGlnLysPheGluAspThrLysPProAlaPheValaLysAlaAspH 460
Db      1377  GTTTCGGGACCCGGCTCAGTGCTTGAAGAACAAGCCGGCTTTGTCAAAAGCCGACCA 1436
Qy      460  sAlaAspGluValaLysPheValaPheGlyGlyAlaPheLeuLysGlyAspLleValaMetP 480
Db      1437  CGCTATGAAGTCCGCTTGTGTGGTGGTGGTCTTCTGAAAGGGGACATGTATATGTT 1496
Qy      480  eGluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyPTrpAlaTh 500
Db      1497  CGAAGGACCCACGAGGAGGAGAAATTAAGTGAAGCCGGAAGATGAATGAATCTGGGCTAC 1556
Qy      500  rPheAlaArgThrGlyAsnProLeuGlyAsnAspLeuSerLeuTyPProAlaTyAsnLe 520
Db      1557  CTTTGCTGAAACCGGAATCTTAATGGGAACGACCTGTCTCTGTCGACGCTTATTAATCT 1616
Qy      520  uThrGluGlnLysLeuGlnLeuAePLeuAeMetSerLeuGlyGlnArgLeuLysGluP 540
Db      1617  GACTGAGAGTACTGACGCTGACTTGAACATGAGCTTCGAGACAGAGACTCAAAAGAAC 1676
Qy      540  cArgValaAspPheTyPThrSerThrLleProLeuLleLeuSerAlaSerAspMetLeuH 560
Db      1677  GCGGGTGAAGTTTGAACAGACACATCCCTCATCTGTCTGCTCCGACATGCTCCA 1736
Qy      560  sSerProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePhePhePheGlyAl 580
Db      1737  CAGGCTCTTTCTTCTTAACCTTCTCTCTCTCTCCAGGCTTTCTTTCTTTTGTGC 1796
Qy      580  aPro 581
Db      1797  TCCT 1800

```

```

RESULT 10
LOCUS   BC069548 1896 bp mRNA linear PRI 30-JUN-2004
DEFINITION Homo sapiens hypothetical protein FLJ31547, mRNA (cDNA clone
ACCESSION BC069548
VERSION   BC069548.1 GI:46854430
KEYWORDS  MGC.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1896)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Topchik, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McKean, P.J., McKernan, K.J., Malek, U.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W., Vittalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, B.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1896)
 Strausberg, R.
 Direct Submission
 Submitted (29-APR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomes Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Baylor Human Genome Sequencing Center
 cDNA Library Preparation: Baylor Human Genome Sequencing Center
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HQSC
 Web site: <http://www.hqsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hultky, S.W., Louesged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRBR Plate: 4 Row: F Column: 1.
 Location/Qualifiers

FEATURES

source

gene

CDS

1..1896
 /gene="FLJ31547"
 /db_xref="LOCUSID:221223"
 120..1697
 /gene="FLJ31547"
 /codon_start=1
 /product="hypothetical protein FLJ31547"
 /protein_id="AAH69548.1"
 /db_xref="GI:46854431"
 /db_xref="LOCUSID:221223"
 /translation="MSGNWVHGOILIMAIWVLAAPTKGPSAEGPDRNTRLGWIOGKOVTVIGSPVAVNVLGVPRAAPLGSIRPTNQPASPMDNLEARSYPLCLQNSEWLLDQHLKHYPKPVSDECLYNTIYAPHADTGSKLPLVWFGAGARTGSASIFDSDALAIVEDLVVVVQIRIGIFGFTTWQHAFGNNAFDQVAALSWVKNIEFPGDDPS

SVTFESAGASIVSSLLSPMAKGLPHKAINESGVALIPYLEANDYEKSEDLQVVAH
FCGNNAASEALRCIARTPKSEKLLTSQKTSFTRVVDGAPFNEPDLISOKAFKA
IPSIIGVNNHCGFLIPMKBAPELISGSKSLALHIONLIHIPOYULVANEYEH
KHSLTRIDSLIDLDGVFVVPVPAITARYHREGATBEKLSRQMKYKATPATGN
PNEGUSLMPATNLTQYIQDLDMNSLQRLKRPVDFTWSTIPILISDSMDLHSPLS
SLTFLSLQPPFFFCAP

ORIGIN

Alignment Scores:

Pred. No.: 6,356-212 Length: 1896
Score: 2606.00 Matches: 503
Percent Similarity: 88.77% Conservative: 3
Best Local Similarity: 88.25% Mismatches: 6
Query Match: 84.64% Indels: 58
DB: Gaps: 2

US-10-023-515-2 (1-581) x BC069548 (1-1896)

12 TrpCyphophleuLeuGlnProleuLeuGlnHlaArgGlnTrpGlyLeuThrGly 31
159 TGGGCTATCTGGGCTCTTGCACGCC-----ACCAAGG 194
32 ProSerAlaGluGlyProGlnArgAnthArgLeuGlyTrpIleGlnGlyLeuVal 51
195 CCTTGTCTGAAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAGCAAGTC 254
52 ThrValLeuGlySerProValProValaenValPheLeuGlyValProPheAlaPro 71
255 ACTGTGCTGGAGAGCCGTGTGCTGTGAACGTGTCCTGGAGTCCCTTCTGCTGCC 314
72 ProleuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 91
315 CCGCTGGATCCCTCGATTCAGAACCCGACGCTGCATCCCTGGGATTAACCTTGGCA 374
92 GluAlaThrSerTrpProAsnLeuGlyLeuGlnAsnSerGluTrpLeuLeuLeuArg 111
375 GAAGCACTCTTAATTTGTCCTCAGAACTCAGAGTGGCTGCTTAAATCA 434
112 HisMetLeuValHisIleTrpProGlyPheGlyValSerGluAspCysLeuTrpLeuAsn 131
435 CACATGCTCAAGGTCAATTAACCGAAATTCGAGGTCTGAGAGCTGCTTCACTGAAC 494
132 IleTrpAlaProAlaHisAlaAspThrGlySerLeuProValLeuValTrpPhePro 151
495 ATCTATGGGCTGCCACCGCATACAGGCTCAAGCTCCCGTCTGGTGTGCTTCCCA 554
152 GluGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleArg 171
555 GAGGTGCTTCAAGACTGGCTCAGCTCATCTTGATGGGTCCGCTGCTGCTCAT 614
172 GluAspValLeuValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrThr 191
615 GAGGACGTGCTGGTGTGCTGCTCAGATACCGCTAGGAATTTGGTTTCTTCAACCA 674
192 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLeuAspGlnValAlaAlaLeuSerTrp 211
675 TGGATACGATGCTCCGGGAACCTGGGCTTCAAGAACACAGGCTGCTGCTGCTGG 734
212 ValGlnLeuAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 231
735 GTCCAGAGAAACATGAGTTCTTGGTGGGACCCAGCTCTGTACCATCTTTGGCGAG 794
232 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLeuGlyLeuPhe 251
795 TCCGGGGAGGCAATAGTGTCTATGCTTATACGTCTCCCATGCGCAAGGTTTATTC 854
252 HisLeuAlaIleMetGluSerGlyValAlaIleIleProGlyLeuGluAlaHisAspTrp 271
855 CACAAAGCATATGAGAGTGGGTGGCTGACATCCCTTACCTGAGGCCCATGATATAT 914
272 GluLeuSerGluAspLeuGlnValValaHisPheCysGlyAsnAsnAlaSerAspSer 291
915 GAGAGAGTGAAGCACTGAGGTGGTGGACATTTCTGTGGTAAACAATGCGTCAAGCTCT 974

292 GluAlaLeuLeuArgCysLeuArgThrIlePheProSerLeuGlyLeuLeuThrLeuSerGln 311
975 GAGGCTCTGCTAGAGTGTGCTGAGCAAAACCTTCAAGAGACTGTGACCTTCAAGCCAG 1034
312 IyethrIySerPheThrArgValValaAspGlyAlaPhePheProAsnGluProLeuAsp 331
1035 AAAACAAAGTCTTTCACCTGAGTGGTGTGATGGTCTTCTTCTTAAATGAGCTCTAGAT 1094
332 LeuLeuSerGlnLeuAlaPheValAlaIleProSerIleIleGlyValaAsnHisGlu 351
1095 CTATGTCTCAGAAAGCAATTAACCAATTCCTTCATCATGTGAGTCAATTAACCAAGAG 1154
352 CysGlyPheLeuLeuProMetLeuGlnAlaProGluIleLeuSerGlySerAsnIySer 371
1155 TGTGGCTTCTGCTGCTGCTTATGAGAGGCTCTTGAATCTTCAAGTGTGCTTCAAGAGTCC 1214
372 LeuAlaLeuHisAlaLeuIleGlnAsnIleLeuHisIleProProGlnTrpIleuHisLeuVal 391
1215 CTGGCTTCCATCTGATCAAAACATCCCTGACATCCGCTCAGATTTGCAACCTTGTG 1274
392 AlaAsnGluTrpPheHisAspIyHisSerLeuThrGluIleArgAspSerLeuLeuAsp 411
1275 GCTAATGAATACTTCCATGACAAAGCACTCCCTGACTGAATCCAGACAGTCTTGTGAC 1334
412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTrpHisAspAsp 431
1335 TTGCTTGAAGATGTCTTCTGTGCTCCTGCACTGAATCAAGCTCGATATACAGA--- 1391
432 AlaGlyAlaProValTrpPheTrpGluPheArgHisArgProGlnCysPheGluAspThr 451
1391 ----- 1391
452 LysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGlyAla 471
1391 ----- 1391
472 PheLeuLeuGlyAspIleValMetPheGluGlyAlaThrGluGluGlyLeuLeuSer 491
1392 -----GAGAGGCAAGAGAGAGAGAGAGATTAAGTCTGAGC 1424
492 ArgLysMetMetLeuTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 511
1425 CGGAAGATGATGAATATCTGGCTACCTTGTGTGACACGGGAATCTTATATGGAACGAC 1484
512 LeuSerLeuTrpProAlaTrpAsnLeuThrGluGlnTrpLeuGlnLeuAspLeuAsnMet 531
1485 CTGTCTGTGGCCAGCTTATATCTGACTGAGCAGTACCTTCAAGCTGACTTGAACATG 1544
532 SerLeuGlyGlnArgLeuLeuGluProArgValaAspPheTrpThrSerThrIleProLeu 551
1545 AGCTCTGAGACAGAGCTCAAAAGAACCGGGGTGATTTTGGACCAAGCAACATCCCTCG 1604
552 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571
1605 ATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCTTAACCTTCTCTCTCTC 1664
572 LeuGlnProPhePhePhePheCysAlaPro 581
1665 CTCCAGCTTCTTCTTCTTCTTGTGCTCTCT 1694

RESULT 11
AX714052 1962 bp DNA linear PAT 15-APR-2003
LOCUS AX714052
DEFINITION Sequence 736 from Patent EP1293569.
ACCESSION AX714052
VERSION AX714052.1 GI:29888980
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1

AUTHORS Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamehika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.

TITLE Full-length cDNA

JOURNAL Patent: EP 1293569-A 736 19-MAR-2003; Helix Research Institute (JP) ; Research Association for Biotechnology (JP)

FEATURES

source location/Qualifiers

1..1962 /organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 6,64e-212 Length: 1962

Score: 2606.00 Matches: 503

Percent Similarity: 88.774 Conservative: 3

Best Local Similarity: 88.254 Mismatches: 6

Query Match: 84.644 Indels: 58

DB: Gaps: 2

US-10-023-515-2 (1-581) x AK714052 (1-1962)

QY 12 TTPCyvRhePheLeuLleuGlnProLeuLeuGlnHleArgGlntrpGlyLeuThrGly 31

DB 189 TGGGCTATCTGGGCTCTTGCAGAGCC-----ACCAAGG 224

QY 32 ProSerAlaGlnGlyProGlnArgAntPheGlnGlyTyrPheGlnGlyVal 51

DB 225 CTTTCTGCGAAGGCGCAAGAGAACACAGCTGGATGATTCAGGCGAAGCTC 284

QY 52 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaPro 71

DB 285 ACTGCGCTGGAGAGCCCTGCTGCTGAACGTCTTCTCGAGTCCCTTGTGCTGCC 344

QY 72 ProLeuGlySerLeuArgPheThrAnProGlnProAlaSerProTyrAspAsnLeuArg 91

DB 345 CCGCTGGGATCCCTCGATTACGAACCGCAGCTGCATGCGCTGGGATACCTTGA 404

QY 92 GlnAlaThrSerTyrProAsnLeuGlnGlnAsnSerGlntrpLeuLeuAspGln 111

DB 405 GAGGCACTCTTACCTTAATTTGCTCCGAACCTAGAGTGTCTTGAATCA 464

QY 112 HisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeuAsn 131

DB 465 CACATGCTCAAGGTGCTTACCCGAAATTCGAGGTGTCAGAAAGCTGCTTACCTGAC 524

QY 132 IleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTyrPhePro 151

DB 525 ATCTATGGCTGCGCCACAGCATACAGGCTCCAAAGCTCCCGCTTGTGCTGCCA 584

QY 152 GlnGlyAlaPheLysThrGlySerAlaSerLlePheAspGlySerAlaLeuAlaTyr 171

DB 585 GAGAGTGGCTTCAAGACAGCTGAGCTCCATCTTGAATGGGTCGCGCTGGCTCAT 644

QY 172 GlnAspValLeuValValValGlnTyrArgLeuGlyLlePheGlyPhePheThr 191

DB 645 GAGACAGTGTGCTGTGTGCTGCTCAAGTACCGGCTAGGAATATTTGGTTCTTACACACA 704

QY 192 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaLeuSerTyr 211

DB 705 TGGGATGAGCATGGCTCGGAGAACTGGGCTTCAAGACCAAGGTGCTGTCTGTGG 764

QY 212 ValGlnLysAsnLleGlnPheGlyGlyAspProSerSerValThrLlePheGlyGlu 231

DB 765 GTCCAGAAAGACATGAGTCTTTCGATGGGAGCCCAAGCTCTGTGACCATTTGGGAG 824

QY 232 SerAlaGlyAlaLleSerValSerSerLeuLleLeuSerProMetAlaLysGlyLeuPhe 251

DB 825 TCGCGGGAGCCATAGTGTTCATAGTCTTAATCTCTCCATGCGCAAGGCTTAATTC 884

QY 252 HisLysAlaLleMetGlnSerGlyValAlaLleIleProTyrLeuGlnAlaHisAspTyr 271

DB 885 CCAAGGCAATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGATTAT 944

QY 272 GlnLysSerGlnAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 291

DB 945 GAAAGAGTGAAGAACTTCAGAGTGGTGGCAATTCGTGTGTAACTAAGCTCAGACTCT 1004

QY 292 GlnAlaLeuLeuAspCysLeuArgThrLysAspSerLysGlnLeuLeuThrLeuSerGln 311

DB 1005 GAGGCGCTGTGAGGTGCTGAGGCAAAACCTTCCAAAGAGCTGTGACCTTACGCCAG 1064

QY 312 LysThrLysSerPheThrArgValAlaAspGlyAlaPhePheProAsnGlnProLeuAsp 331

DB 1065 AAAACAAAGTCTTTCACCTCGAGGTGTGATGTGCTTCTTCTTAATGAGCCTAGAT 1124

QY 332 LeuLeuSerGlnLysAlaPheLysAlaLleProSerLleIleGlyValAsnAsnHisGln 351

DB 1125 CTATTGTCTCGAAAGCAATTTAAAGCAATTCCTTCCATTAACGAGTCAATACACAGAG 1184

QY 352 CysGlyPheLeuLeuAspProMetLysGlnAlaProGlnLleLeuSerGlySerAsnLysSer 371

DB 1185 TGTGGCTTCTGTGCTGCTTATGAAGAGGCTCTGAGATCTCAGTGGCTCCAAACAATCC 1244

QY 372 LeuAlaLeuHisLleuLleGlnAsnLleLeuHisGlnProGlnTyrLeuHisLysVal 391

DB 1245 CTTCGCTTCCATGTATACAAACATCTGCACATCCGCTCAGATTGTCACCTTGTG 1304

QY 392 AlaAsnGlnTyrPheHisAspLysHisSerLeuThrGlnLleArgAspSerLeuLeuAsp 411

DB 1305 GCTAATGATATCTTCATGACAAAGCACTCCCTGACTGAATCCGAGACATCTTGTGAG 1364

QY 412 LeuLeuGlnLysPheValPhePheValValProAlaLeuLleThrAlaArgTyrHisAspAsp 431

DB 1365 TTGCTTGAAGATGTGTTCTTGTGTGCTCCGCACTGATCAAGCTGATATCAGAGA--- 1421

QY 432 AlaGlyAlaProValTyrPheTyrGlnPheArgHisArgProGlnCysPheGlnAspThr 451

DB 1421 ----- 1421

QY 452 LysProAlaPheValLysAlaAspHisAlaAspGlnValArgPheValPheGlyAla 471

DB 1421 ----- 1421

QY 472 PheLeuLysGlyAspLleValMetPheGlnGlyAlaThrGlnGlnLysLeuLeuSer 491

DB 1422 -----GAAAGGCCACGAGAGGAGAGAAATTACTGAGC 1454

QY 492 ArgLysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 511

DB 1455 CGGAAGATGATGAATACTGGGCTACCTTGTCTCGAACCGGGAATCTTAATGGAAAGCA 1514

QY 512 LeuSerLeuTrpProAlaTyrAsnLeuThrGlnGlnTyrLeuGlnLeuAspLeuAsnMet 531

DB 1515 CTGTCTCTGTGGCACCTTAATATCTGACGTGAGCTACCTCAGCTGACTTGAACATG 1574

QY 532 SerLeuGlnGlnArgLeuLysGlnProArgValAspPheThrThrSerThrLleProLeu 551

DB 1575 AGCTTCGACAGAGACTCAAAAGAACCGCGGTGGATTGGAACCAAGCAATCCCCCTG 1634

QY 552 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571

DB 1635 ATCTGTGTGCTCGACATGTCTCCACAGTCTCTTCTTCTTAACCTTCTCTCTGC 1694

QY 572 LeuGlnProPhePhePheCysAlaPro 581

DB 1695 CTCACGCTTCTTCTTCTTGTGCTCTCT 1724

RESULT 12

AK056109 1962 bp mRNA linear PRI 30-JAN-2004

LOCUS Homo sapiens cDNA FLJ31547 f1e, clone NT2R12001010, weakly similar

DEFINITION to PATTY ACYL-CoA HYDROLASE PRBCURSOR, MEDIUM CHAIN (EC 3.1.2.14).

ACCESSION AK056109
VERSION AK056109.1 GI:16551422
KEYWORDS oligo capping; file (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, R., Sekine, M., Ohsashi, M., Nishii, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahata, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, T., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimura, M., Watanabe, M., Hiraoaka, S., Chiba, Y., Iehida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Houchi, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsumawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Sato, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kunigai, A., Itakura, S., Fukuzumi, T., Fujimori, Y., Komiyama, M., Tashiro, H., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Ozaki, K., Hirose, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, H., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Maeuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
2
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahata, K., Masuho, Y., Nagai, K. and Isegai, T. NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 1662)
Isegai, T., Otsuki, T. and Sugiyama, T. Direct Submission
Submitted (24-OCT-2001) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5- & 3-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES
Source location/Qualifiers
1..1662
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2R12001010"

CDS
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2R12"
/note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction. -majorly NT2 neuron"
150..1727
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71094.1"
/db_xref="GI:16551423"
translation="MSGMWVHGGQILIAWIAPTKPSAEGPQNRTRLGIQKQ
/TVLAGSPVAVVFLGVPAPADPLSGIRFTNPQSPASPMNDLRATSYPMICONSFML
LDQHLKHYPRKFGVSDCLYNTIYAPADPTGSLPLVWPGGAPKTSASIFDGS
ALAAEDLVVVOYRGLGIPGFTTWDQAHGNAFKOVAALSMVKNIIEFGDPS
SVTFGEBSAGAISVSSLSLSPMAKGLFKALMESVALITPLIANDYEKSELOVYAH
FCGNNAISSEALRLRLTRKPESEKELTTSQKTSFTRVVDGAFEPNEPLDLSQKAFKA
IPSIIGVNNHSCGFLPMKEBAPETLSGNKSLALHILIONLHIIPQYHLVANFEHD
KSLITEIRDSLLDLGDVFEVFPALITARYHREGATEEKLISRMQKYMATFATGN
PNGNDLSIMPVNNLTQYLOLDLMSLGQRLEKPRVDWPTSTIPLISADWLHSPLS
SLTFSLIQPPFPFCAP"
ORIGIN
Alignment Scores:
Pred. No.: 6,646-212 Length: 1962
Score: 2606.00 Matches: 503
Percent Similarity: 88.77% Conservative: 3
Best Local Similarity: 88.25% Mismatches: 6
Query Match: 84.64% Indels: 58
Gaps: 9
US-10-023-515-2 (1-581) x AK056109 (1-1962)
QY 12 TTPCySPhePheLeuIleLeuGlnProLeuLeuGlyHisArgGlnTTPGlylySthrgly 31
189 TGGGCTATCTGGTCTGCTTGCAGCCCC-----ACCAAGGG 224
QY 32 ProSerAlaGluGlyProGlnArgGlnThrArgLeuGlyTTPILGnglyGlnVal 51
225 CCTTGCTGAAGGGCCACAGAGAACCCAGCTGGAGTTCAGGGCAGCAGATC 284
QY 52 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIaPro 71
285 ACTGTCCTGGAGAGCCCTGCTGCTGAACGTTCTCCGAGAGTCCCTTCTCTCTCC 344
QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeuArg 91
345 CCGCTGGGATCCCTCGATTTCAGAACCCGACGCTGCATCCCTGGGATTACTTGCAG 404
QY 92 GluAlaThrSerTyrProAsnLeuLeuGlnAsnSerGlnTTPLeuLeuAsnGln 111
405 GAAGCACCTCTTACCTTAATTGCTGCCAGAACTCAGATGAGTCTGCTTAAGTCAA 464
QY 112 HisMetLeuValValHisIleTyrProLysPheGlyValSerGluAspCysLeuTyrLeuAsn 131
465 CACATGCTCAAGTGCATTAACCCGAATTCCGAGTGTGAAGAATGCTCTTACTGTAAC 524
QY 132 IleTyrAlaProAlaHisAlaAspThrGlySerIlePheAspGlySerAlaLeuAlaIaTyr 151
525 ATCTATGGCCGCTCCACGCGATTCAGAGCTTCAAGCTCCCGTCTTGTTGGTTCCTCA 584
QY 152 GlyGlyAlaPheIleValThyGlySerAlaSerIlePheAspGlySerAlaLeuAlaIaTyr 171
585 GAGGTGCTTCAACACGCGCTCCACCTTCATCTTGAAGGATCCGCTGCTGCTCTCT 644
QY 172 GluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 191
645 GAGGACGTGCTGCTGTTGTGTCTCAGTACCGGCTAGGAATATTGTTGTTCTTCCACA 704
QY 192 TTPAspGlnHisAlaIaProGlyAsnTTPAlaPheLysAspGlnValAlaAlaLeuSerTTP 211
705 TGGATTCAGCATGCTCCGGGGGAACTTCAAGGACCAAGTGTGCTCTGCTCTGCG 764

QY 212 ValGlnLysAsnIleGluPheGlyGlyAspProSerSerValThrIlePheGlyGln 231
DB 765 GTCAGAGAAATCGAGTCTTCGGTGGGAGCCCAAGCTCTGTACCATCTTTGGCGAG 824
QY 232 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeuPhe 251
DB 825 TCCGCGGAGCCCAATAGTGTCTTACGTATCTGTCTCCCAATGGCCAAAGGCTTATTC 884
QY 252 HisLysAlaIleMetGluSerGlyValAlaIleIleProGlyLeuGlnAlaHisAspTyr 271
DB 885 CACAAAGCCATCATGAGAGAGTGGGGTGGCATATCCCTTACCTGGAGGCCCATGTATAT 944
QY 272 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 291
DB 945 GAGAAAGTAGAGGACCTCGACAGTGGTGGCACTTTCTGTGTAAACAATGGTCAGACTCT 1004
QY 292 GluAlaLeuLeuAspGlyCysLeuAspGlyThrIleProSerLysGlyLeuLeuThrLeuSerGln 311
DB 1005 GAGGCGCTCTGAGGCTGAGGAGCAAAACCTCCAGAGAGCTGCTGACCTCGAGCCAG 1064
QY 312 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 331
DB 1065 AAAACAAAGCTCTTTCATCGAGAGCTTGAATGGTCTTTCTTCTTATAGAGCTCTAGAT 1124
QY 332 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHisGln 351
DB 1125 CTATTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCATTATCGAGTCAATACCAAGAG 1184
QY 352 CysGlyPheLeuLeuProMetLysGlnAlaProGlnIleLeuSerGlySerAsnLysSer 371
DB 1185 TGTGCTTCTCTGCTCTGATGAAGAGGCTCTGAGATCTCAAGTGGCTCCAAAGTCC 1244
QY 372 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleLeuHisLeuVal 391
DB 1245 CTGGCCCTCGCATGTATCAAAACATCCCGACATCCGCGCTCAATATTGACCTTGTG 1304
QY 392 AlaAsnGluTyrPheHisAspLysHisSerLeuThrGlnIleArgAspSerLeuLeuAsp 411
DB 1305 GGTAAATGAATCTTCATGACAAAGCATCCCTGACTGAATCCGAGACAGTCTTCTGAGC 1364
QY 412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAsp 431
DB 1365 TTGCTGGAGATGCTTCTTGTGTGCTCCCTGACATGATCAAGCTCGATATCAAGAG--- 1421
QY 432 AlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAspThr 451
DB 1421 ----- 1421
QY 452 LysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGlyAla 471
DB 1421 ----- 1421
QY 472 PheLeuLysGlyAspAlaValMetPheGluGlyAlaThrGluGluGlnLysLeuLeuSer 491
DB 1422 -----GAGAGAGCCAGAGAGAGAGAAATTACTGAGC 1454
QY 492 ArgLysMetMetLysTyrTTPAlaIleThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 511
DB 1455 CGAAAGATGATGAATACTGGGCTAACCTTGTCTGAAACCGGGAATCCTTAATGGGAAGAC 1514
QY 512 LeuSerLeuTTPProAlaTyrAsnLeuThrGluGlnIleLeuGlnLysAspLeuAsnMet 531
DB 1515 CTGCTCTGTGGCCAGCTTATATCTGACTGACGATACCTCCAGCTGGAGCTTGAACATG 1574
QY 532 SerLeuGlnAlaGlyLeuLysGluProArgValAspPheThrPheSerThrIleProLeu 551
DB 1575 AGCTTCGAGACGAGACTCAAGAAACCGGGGTGATTTTGGACCAACATCCCTCCCTG 1634
QY 552 IleLeuSerLysSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571
DB 1635 ATCTGTCTGCTCTCGACATGCTCCACAGTCTCTTCTTCTTAACTTCTCTCTCTC 1694
QY 572 LeuGlnProPhePhePhePheCysAlaPro 581

DB 1695 CTCAGAGCTTCTTCTTCTTGTGCTCT 1724
RESULT 13
AX958392
LOCUS
DEFINITION
SEQUENCE 20 from Patent WO0226588.
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1857
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,91e-210 Length: 1857
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 84.07% Indels: 23
DB: 6 Gaps: 1
US-10-023-515-2 (1-581) x AX958392 (1-1857)
QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTTPriIleGlnGlyLysGln 50
DB 274 GGGCTTCTCTGTAAGGCGCACAGAGAACACAGCTGGAGATGATTCAGGGCGAGCA 333
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 70
DB 334 GTCACTGTCTGGAGAGCCCTGTGCTGAGAGTGTCTCGAGTCCCTTGTGCTGCT 393
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeu 90
DB 394 CCCCCGTGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGGATTACTTG 453
QY 91 ArgGluAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTTPLeuLeuAsp 110
DB 454 CGAAGAGCCACTCTTAACCTTAATTTGTGCTCCAGAACTCAAGATGGCTGCTTACAT 513
QY 111 GluHisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeu 130
DB 514 CAACATATGCTCAAGGTGATTAACCCGAAATTCGAGGTGTCAGAACTGCTTACCTG 573
QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTTPPhe 150
DB 574 AATCATTAAGCGGCTGCCACCGCGGATACAGGCTCCAACTCCCTCTTGGTGTGCTC 633
QY 151 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 170
DB 634 CCAAGAGGTGCTTCAAGACTGAGCTCAGCTCATCTTTGATGGGTCGCGCTGGCTGCG 693
QY 171 TyrGluAspValLeuValValAlaGlnTyrArgLeuGlyTlePheGlyPhePheThr 190
DB 694 TATAGAGAGTGTGTTGTGTGCTGCTCAAGTACCGGCTAAGAAATTTGGTTCCTTCAAC 753
QY 191 ThrTTPAspGlnHisAlaProGlyAsnTTPAlaPheLysAspGlnValAlaAlaLeuSer 210
DB 754 ACATGGGATCAAGATGCTCCGGGGAAGTGGGCTTCAAGAGACCAAGTGGCTCTGTGCC 813
QY 211 TTPValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230

Db 814 TGGGTCAGAGAAACATCGAGTTCTTCGGTGGGAGCCCGAGCTGTGACCATCTTTGGC 873
Qy 231 GUSERRAAGLYAAlIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu 250
Db 874 GAGTCCGGGGGAGCCATAGTGTCTTACTTAACTGTCCTCCATGGCCAAAGGCTTA 933
Qy 251 PheHISLYAAlIleMetGlySerGlyValAlaIleIleProTyrLeuGlyAlaHISaP 270
Db 934 TTCCCAAGGACCATCATGAGAGAGTGGGGGCGGCATCATCTCTTACCTGAGGCCCCATGAT 993
Qy 271 TTTGlyLeuSerGlyLeuSerGlyLeuValAlaAlaHISpHeCyGlyAAsnaanaIaSerAaP 290
Db 994 TATGAG 1053
Qy 291 SerGlyAlaIleLeuLeuArgCyAlaLeuArgTyrIleProSerIleIleGlyValaAsnaHIS 310
Db 1054 TCTGAGGCGCTTCCTAG 1113
Qy 311 GlnLYSThrIleSerPheThrArgValValaAspGlyAlaPhePheProAsnGlyProLeu 330
Db 1114 CAGAAACAAAGCTTTTCACTCGAGTGGTGAATGGTCTTTCTTCTTAATAGGCTCTTA 1173
Qy 331 AspLeuLeuSerGlyLeuValaPheLeuValaIleProSerIleIleGlyValaAsnaHIS 350
Db 1174 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATAACCA 1233
Qy 351 GlnCYGlyPheLeuLeuProMetIysGlyAlaProGlyIleLeuSerGlySerHISLY 370
Db 1234 GAGTGTGGCTTCCTCTGCTGCTAGT----- 1257
Qy 371 SerLeuAlaLeuHISLeuIleGlnHISLeuHISLeuProGlnTyrIleuHISLeu 390
Db 1258 -----CACATCCCGCTCAAGATTTGACCTT 1284
Qy 391 ValAlaAsnGlyTyrPheHISaPlyHISaPlyHISaPlyHISaPlyHISaPlyHISaPly 410
Db 1285 GTGGGTAAATGAACTTTCATCATGACAAAGCACTCCCTGACATCGAAATCGAGAGAGTCTTCG 1344
Qy 411 AspLeuLeuGlyAAsnaHISaPlyPhePheValValProAlaLeuIleThrAlaArgTyrHISaP 430
Db 1345 GACTTGCTGGAGAGTGTCTTTGTGGTCCCTGACATGATCAAGCTCATATACACAGA 1404
Qy 431 AspAlaGlyAlaProValTyrPheTyrGlyPheArgHISaPlyProGlnCYGlySerHISaP 450
Db 1405 GATGCTGGGACCTGTCTACTTCTATGAGTTTGGACCGGCTCAAGTCTTGAAGAC 1464
Qy 451 ThrLYSProAlaPheValIleValaAspHISaPlyHISaPlyHISaPlyHISaPlyHISaPly 470
Db 1465 ACGAAGCCAGCTTTGTCTCAAAAGCCAGCAAGCTGATGAAAGTCCGCTTGTGTGGTGT 1524
Qy 471 AlaPheLeuLYSGLYAspIleValMetPheGlnGlyAlaThrGlnGlnGlnGlnGlnGln 490
Db 1525 GCCTTCTCGAAGGGGAGCATTTGTATGTTTGAAGAGGACCGAGAGAGAGAGAGAGAGAG 1584
Qy 491 SerArgLYSMeMetLYSTyrTTPAlaThrPheAlaArgThrGlyAAsnaPheAsnGlyAAsn 510
Db 1585 AGCCCGAAGATATGAAATATCTGGGCTACCTTGTCTGAACCGGGAATCTTAATGGGAAC 1644
Qy 511 AspLeuSerLeuTyrProAlaTyrAsnLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGln 530
Db 1645 GACCTGTCTCTGTGGCCAGCTTAATCTGATGAGAGATCACTCCAGCTGAGCTTGAAGAC 1704
Qy 531 MetSerLeuGlyGlnArgLeuLYSGLYProArgValaAspPheTyrThrSerThrIlePro 550
Db 1705 ATGAGAGCTCGACAGAGACTCAAGAAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1764
RESULT 14
CQ733044 1716 bp DNA linear PAT 03-FBB-2004
LOCUS CQ733044
DEFINITION Sequence 18978 from Patent WO02068579.
ACCESSION CQ733044
VERSION CQ733044.1 GI:42314969

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
TITLE Kite, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 18978 06-SEP-2002;
FEATURES
source location/Qualifiers
1..1716
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2,2e-203 Length: 1716
Score: 2505.00 Matches: 478
Percent Similarity: 93.98% Conservative: 6
Best Local Similarity: 92.82% Mismatches: 15
Query Match: 81.36% Indels: 16
DB: 6 Gaps: 1
US-10-023-515-2 (1-581) x CQ733044 (1-1716)
Qy 31 GTPROSERALAGLYProGlnArgAsnThrArgLeuGlyTyrIleGlnGlyLYSGln 50
Db 28 GGGCTTCTCGTGAAGGGGCAAGAGAAACACAGAGCTGGATGATCAGGGCAAGCA 87
Qy 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 70
Db 88 GTCACTGTCTGGGAAGGCTGTGCTGGAACGTGTCTTGAGAGTCCCTTGTGCT 147
Qy 71 PROPOLEUGLYSerLeuArgPheThrAsnProGlnProAlaSerProTyrAsnLeu 90
Db 148 CCCCCTGGAGTCCCTGCAATTTACGAACCCGAGCTGCAATCGCCCTGGATTAATCTTG 207
Qy 91 ARGGLUALATHSerTyrProAsnLeuCYSGlnAsnSerGlyTyrLeuLeuAsp 110
Db 208 CGAAGAGCACTCTTACCTTAATTTTGCTCCAGAACTCAGAGTGGCTCTTAAT 267
Qy 111 GlnHISMeLeuLYSValHISArgTyrProLYSPheGlyValSerGlyAspCYSGlyTyrLeu 130
Db 268 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGAGTGCAGAAAGACTGCTTAACCTG 327
Qy 131 AsnIleTyrAlaProAlaHISaPlyHISaPlyHISaPlyHISaPlyHISaPlyHISaPly 150
Db 328 AACATCTATGCGCTGCCCAAGCCGATACAGCTCAAGCTCCCGTCTTGGTGGTTC 387
Qy 151 PROGLYGLYAlaPheLYSThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 170
Db 388 CCAGAGAGGCTTCAAGAGTGGCTCAGCTTCATCTTGAAGGGATCCGCTGGCTGCC 447
Qy 171 TTTGlyAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
Db 448 TATGAGAGAGTGTGCTGGTGTGCTGCTCAAGTACCGGCTTAAGAAATTTGGTTCCTTCAAC 507
Qy 191 ThrTPRAsnGlnHISaPlyProGlyAsnTyrAlaPheLYSAspGlyValAlaAlaLeuSer 210
Db 508 ACATGGAGATCAGAGAGTCCCGGGAGACTGGGCTTCAAGAGACAGAGTGGCTCTGTGC 567
Qy 211 TTPValGlnLYSAsnIleGlyPhePheGlyGlyAAspProSerSerValThrIlePheGly 230
Db 568 TGGTCCAGAGAGAAACATGAGTCTTGGTGGAGACCCAGGCTGTGACCATCTTTGGC 627
Qy 231 GUSERRAAGLYAAlIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu 250
Db 628 GAGTCCGGGGGAGCATAGTGTCTTACTGTTTAACTGTCCTCCATGAGCCAAAGGCTTA 687


```

Qy 251 PheHisIySaIaIleWecGluSerGlyValaIaIleIleProTyrLeuGluIaHisAsp 270
Db 668 TTCCACAAAGCCATCATGAGAGTGGGGTGGGCATCATCTTACCTGAGGCCCATGAT 747
Qy 271 TyrGluYSerSerGluAspLeuGluValaIaHisIspheCyseGlyAAsnaaIaSerAsp 290
Db 748 TATGAGAAAGAGTGAAGACCTGCAGGTGGTTGACATTTCTGTGTAAACAATCGTCAGAC 807
Qy 291 SerGluIaIleuLeuArgCyseLeuArgThryIspProSerIyGluLeuLeuThryIseu 310
Db 808 TCTGAGGCCCTGTCTAGGTGCTGAGAGCAAAACCTCCAAAGACCTGCTGACCTCCAGC 867
Qy 311 GluYThrIySerSerPheThzArgValaIaAspGlyAAlaPhePheProAsnIuProIeu 330
Db 868 CAGAAACAAAGTCTTTCACCTGAGTGTGATGGTGTCTTCTTCTTAATAGCTCTTA 927
Qy 331 AspLeuLeuSerGluIyAlaPheIySaIaIleProSerIleIleGlyValaAsnaaHis 350
Db 928 GATCATATTGTCAGAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATAACAC 987
Qy 351 GluCyseGlyPheLeuLeuProMetIySeGluIaIaProGluIleuSerGlySerAsnIy 370
Db 988 GAGTGTGCTTCTGTCTGTCTGTAGAGAGGCTCTGAGATCTCAGTGGCTCCAAACAG 1047
Qy 371 SerIeuAlaIeUhiIeUleIegIaAsnIleUhiIeIleProProGluIuThryIseu 390
Db 1048 TCCCTGGCCCTCCATCATGATCAAAACATCCCGACATCCCGCTCAGATTTGCACTT 1107
Qy 391 ValaIaAsnGluTyrPheHisAspIyHisIserLeuThrgluIaArgAspSerLeuLeu 410
Db 1108 GTGGTAATGAAATCTCTCCATGACAGACCTCCCTGACATCGAATCCGAGACGTCCTG 1167
Qy 411 AspLeuLeuGluIyAspValaIaPhePheValaIaProAlaIeUleIeThAlaArgTyrHisArg 430
Db 1168 GACTTGCTTGGAGATGTCTTCTTGTGGTCCCTGACATGATCACAGCTGATATACAGA 1227
Qy 431 AspAlaGlyAlaProValaIyTyrPheTyrGluPheArgHisArgProGluIuCyPheGluAsp 450
Db 1228 GATGCTGGTGACCGTGTCTACTTCTATGAGTTTGGGACCGGCTCAGATGCTTGAAGAC 1287
Qy 451 ThrIySProAlaPheValaIyAlaAspHisAlaAspGluValaIaArgPheValaIaPheGly 470
Db 1288 ACGAAGCGGCTTTTGTCAAAAGCCACACCGTGTGTGAAGTCCGCTTGTGTGCGTGTG 1347
Qy 471 AlaPheLeuIySgIyAspIleValaIaMerPheGluGluIyAlaIaThngluIuLeuLeu 490
Db 1348 GCCTTCTGAAAGGGGACATGTTATGTTCCAAAGAGCCACGAGAGAGAAAGTTACTG 1407
Qy 491 SerArgIySMeMetIyTyrTyrAlaIaThzPheAlaArgThrgIyAsnProAsnGlyAsn 510
Db 1408 AGCGGAAAGATGAGAAATATCTGGGCTACTTGTGCTGAAACCGGACG----- 1455
Qy 511 AspLeuSerLeuTyrProAlaIyTyrAsnLeuThrgluIuThryIleuGluIeUAspLeuAsn 530
Db 1456 -----CACTTCTATACAGAGAGGTGTCT 1479
Qy 531 MetSerLeuGluIyArgLeuLeuGluProArgValaAspPheTyr 545
Db 1480 GTGATGATGATCCCAAAATGCAAGTCCCAACCCCAACCCCACTTGG 1524

```

```

RESULT 15
LOCUS AB186392 1728 bp mRNA linear MAM 06-AUG-2004
DEFINITION Canis familiaris cauxin mRNA for carboxylesterase-like urinary
ACCESSION AB186392
VERSION AB186392.1 GI:51014274
KEYWORDS
SOURCE
ORGANISM
Canis familiaris (dog)
Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Plissipedia; Canidae; Canis.
REFERENCE
1

```

AUTHORS Miyazaki, M., Yamashita, T., Talra, H. and Suzuki, A.
 TITLE Published family protein
 JOURNAL Published only in Database (2004)
 REFERENCE 2 (bases 1 to 1728)
 AUTHORS Miyazaki, M., Yamashita, T., Talra, H. and Suzuki, A.
 TITLE Direct Submission
 JOURNAL Submitted (04-AUG-2004) Masao Miyazaki, RIKEN, Frontier System Research, Supra-Biomedical System Research Group, Hiroseawa 2-1, Wako-shi, Saitama 351-0198, Japan
 (E-mail: mmiyazaki@postman.riken.jp, Tel: 81-48-467-9619, Fax: 81-48-467-9620)

FEATURES
 source
 1..1728
 Location/Qualifiers
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /chromosome="2"
 1..1728
 /gene="cauxin"
 1..1728
 /gene="cauxin"
 /codon_start=1
 /product="carboxylesterase-like urinary excreted protein"
 /protein_id="BAD35015.1"
 /db_xref="GI:51014275"
 /translation="MSGEMGHIGQTLIWAIVTVLAAAREGPAADAPVSTRLGWKCKO
 ATVAGSTMPVNVPLGITPAAPRPGRLPRFKRPRKALLIMNDRDSTSPKLCIQNSWML
 SDQHLKVRHPNENVEDCLYLNTYAPAHNATSGKLPVWVWPFEGAFETASAFDS
 ALAAVEEVLLVTQYRLGIFGFRGTQQAAGWAFLDQALATWQENTLFFGDDH
 SVTIFESAGAILSVSGIYVLSPPMAAGLFHKAIMSGVALIFPLRAPDERMEDIQVIR
 ICGCNVDSVALLOCLRAKSEELIDINKTKSFTRVVDGFPFDEPLDLTEKTFNS
 IPSVIGVNHCEGFLIPMKEPRIILGSSNSLHLIHRVLIHINQVLYVADQYFN
 KHSRVERIRDSFLDLIGVRFVVGQYARVHRPAGAPVPEYEQHPEDCKOTRRPV
 KADHSRIRVFEGAFILKNITVMEGATSEKLSRPMRYWAFATGVPNGELPL
 WPAVSQSEQYLKDLINTISVGQKKEQVEPWSDTLPILIMSMSTAPPGPVPPLLSVL
 LPLFSSAP"

CDS

gene

1..1728
 /gene="cauxin"

ORIGIN

Alignment Scores:

Pred. No.: 1.13e-181 Length: 1728
 Score: 2250.00 Matches: 417
 Percent Simlarity: 86.75% Conservative: 61
 Best Local Simlarity: 75.68% Mismatches: 73
 Query Match: 73.08% Indels: 0
 DB: 4 Gaps: 0

US-10-023-515-2 (1-581) x AB186392 (1-1728)

```

Qy 31 GlyProSerAlaGluGlyProGluIaArgAsnThzArgLeuGlyTyrPleIeGluIyLeGln 50
Db 73 GGTCCAGCTGCTGATGACACCGGTGAGAGACCAAGCTGGGATGGCTCCGGGCAAGCA 132
Qy 51 ValThrValleuGlySerProValaIaAsnValaIaPheLeuGlyValaIaProPheAlaIa 70
Db 133 GCCACTGTGCTGGGAGACCAACATGCTGTGAACGTGTCTCTCGGAGATCCCTTGTGTCG 192
Qy 71 ProProIeUuIySerIeUaArgPheThzAsnProGluIuProAlaSerProTyrAspAsnLeu 90
Db 193 CCCCCTTAAAGACCCCTGCGCATTTAAGAGACCAAGCTTCATTTGCTCTGAAATGATCA 252
Qy 91 ArgGluAlaThzSerTyrProAsnLeuCyseLeuGluIaAsnSerGluTyrPleuLeuAsp 110
Db 253 CGAGATCCCAAGCTCTTAATCTTAATTTGTGCTCCAGAACTCAAGTGGCTGCTCAGAT 312
Qy 111 GluHisMetLeuIyValaHisTyrProIySPheGlyValaSerGluIuAspCyseIyTyrLeu 130
Db 313 CAACACTTCTCAAGGTGATTAACCCCAATTTAGAAAGTCCGAAAGACTGCTGTACTT 372
Qy 131 AsnIleTyrAlaProAlaHisAlaAspThzGlySerIyIseuProValleuValaIaTyrPhe 150
Db 373 AACATTCACCAACAGCCACGCGAAACACCGGCTCCCAAGCTCCCGTAAATGATGTGCTC 432
Qy 151 ProGluIyAlaPheIyThrgIySerAlaSerIlePheAspGlySerAlaIaIa 170

```


this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 15, 2005, 09:44:01 ; Search time 5241 Seconds
(without alignments)
4219.680 Million cell updates/sec

Title: US-10-023-515-2
Perfect score: 3079
Sequence: 1 MPOGLTASQWCFLLIQP.....PLSLTFLSLQPPFPFCAP 581

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_spool/6664091/runat_14062005_133641_10470/app_query.fasta.1.775
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALL=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEXRSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=6664091@cgn.1.1_3437@runat_14062005_133641_10470 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hic: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282.5	41.7	2687	3	AK077248 Mus muscu
2	1272.5	41.3	2661	3	AK033563 Mus muscu
3	1264	41.1	1680	9	AY414461 Mus muscu
4	1236	40.1	1958	3	AK078953 Mus muscu
5	1230	39.9	1680	9	AY414459 Mus muscu
6	1197	38.9	4927	3	AK040349 Mus muscu
7	1191	38.7	2888	3	HSM803013 Mus muscu
8	1182	38.4	3909	3	HSM806270 Mus muscu
9	1158.5	37.6	2038	3	BC019926 Mus muscu

10	1142.5	37.1	2356	3	AK037191 Mus muscu
11	1129.5	36.7	1944	3	AK078879 Mus muscu
12	1097.5	35.6	1918	7	CF111083 Mus muscu
13	1069.5	34.7	1826	3	CR618303 full-length
14	1055	34.3	1859	3	CR641368 Tetradon
15	1048	34.0	1817	3	CR634885 Tetradon
16	1038	33.7	1849	3	CR647187 Tetradon
17	1028	33.4	1823	3	CR640788 Tetradon
18	1026	33.3	1853	3	CR723405 Tetradon
19	1020	33.1	1857	3	CR637401 Tetradon
20	895.5	29.1	742	4	B1822069 Mus muscu
21	815	26.5	1610	3	AK050444 Mus muscu
22	804	26.1	1438	7	CF111142 Shultzom
23	804	26.1	1438	7	CF111143 Shultzom
24	780	25.3	4319	3	BC069040 Mus muscu
25	779	25.3	3853	3	AK039018 Mus muscu
26	768	24.9	2547	9	AY402702 Mus muscu
27	763	24.8	2173	3	AK046080 Mus muscu
28	763	24.8	2181	3	AK043748 Mus muscu
29	763	24.8	2184	3	AK042335 Mus muscu
30	760.5	24.7	2187	3	BC026315 Mus muscu
31	760	24.7	2026	3	AK052200 Mus muscu
32	755.5	24.5	1723	9	AY407736 Homo sapi
33	755.5	24.5	1723	9	AY407737 Pan trogl
34	754	24.5	1723	9	AY407738 Mus muscu
35	752	24.4	1810	3	CR611139 full-length
36	752	24.4	2177	3	AK036259 Mus muscu
37	752	24.4	2523	6	CD014089 Mus muscu
38	750.5	24.4	576	4	B1824830 Mus muscu
39	746	24.2	3206	3	AK050337 Mus muscu
40	737	23.9	40	9	AY402703 Pan trogl
41	733.5	23.8	1113	7	CN066114 F05 Ag2_P
42	726	23.6	1745	3	BC036813 Homo sapi
43	718.5	23.3	1117	7	CN061747 Homo sapi
44	717	23.3	818	7	CN067323 J05_Ag2_P
45	716	23.3	3431	3	BC041823 Homo sapi

ALIGNMENTS

RESULT 1
AK077248
LOCUS
DEFINITION
Mus musculus 11 days pregnant adult female ovary and uterus cDNA.
Riken full-length enriched library, clone:5031415B19
Product: similar to CARBOXYL ESTERASE PRECURSOR (EC 3.1.1.1)
(Alc-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE ESTERASE)
(PROCAINE ESTERASE) (METHYLUTRYLASE) [Mesocricetus auratus], full
insert sequence.

ACCESSION
AK077248
VERSION
AK077248.1 GI:26346111
KEYWORDS
HTC; Cap trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Carninci, P., and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumaki, T., Tashiro, H., Itoh, M.,
3

LOCUS	DEFINITION	AK033563	2661 bp	mRNA	linear	HTC 03-APR-2004
Qy	270	AerpTyrGluYvSserGlnApsLeuAlaValAlaValIlePheCysGlyYAsnAAspLeuSer	289			
Db	848	TCCTCTGAAATGGTTACAC-----AGAAATAGTGGCAAACTTACTCTGGCTGTGCACCTGTG	901			
Qy	290	AAspSerGlnAlaLeuLeuAArgCysLeuAArgThrIvsProSerIvsGluLeuLeuThrIleu	309			
Db	902	AACCTCAGAGACCCCTGATGTGGCTGCTTAAGAGCAAAAGATGAAGCAGAGATGCTGGCTATT	961			
Qy	310	SerGlnIvsThrIvsSerPheThrArgValYAlaAspGlyAlaPhePheProAsnGluPro	329			
Db	962	AATAAGGCTTCAAAATCATCTCCCTGGTGTGGATGAGATGAGATTCCTACCAAGCATCT	1021			
Qy	330	LeuApsLeuLeuSerGlnIvsAlaAAspLeuAlaIleProSerIleIleGlyValAAsn	349			
Db	1022	CAGAGCTGAAGGCTCTTAAGGATTTTACCCCTGTCCAGCATCATTTGTATCAACAT	1081			
Qy	350	HisGluCysGlyPheLeuLeuPro-----	357			
Db	1082	GATGAATATGGTTGGATTTCTTCTTAAGCATCATGACCCAGCTCAGAAATAGAGGAATA	1141			
Qy	358	---MecIvsGluAlaProGluIleLeuSerGlySerAAsnIvsSerLeuAlaLeuHileu	376			
Db	1142	ACCAAGAAAGACCTGCACGCTGTCTTA-----AAAGACAGCCCTAAATAATG	1189			
Qy	377	IleGlnAAsnIleLeuHileProProGlnIvsIleuHileuValAlaAAsnGluIvsPhe	396			
Db	1190	ATG-----CTGCCTCTCAAGTGTGTGAACCTGATTAAGAGAGATGACATG	1234			
Qy	397	HisApsIvsHisSerLeuThrGluIleAArgAAspSerLeuAAspLeuLeuGlyAAspVal	416			
Db	1235	GGGACACATGACAGACCCAGACACTGTGCACACAGTTCAAGAGATGAAGGGGACCTTC	1294			
Qy	417	PhePheValValProAlaLeuIleThrIlaAArgIvsHisAArgAAspAlaGlyAlaProVal	436			
Db	1295	ATGTTGCGTATCCCTGCACATCAAGTACACATTTTCAACG---TCCCATGCTCTGTG	1351			
Qy	437	TyrPheIvsTyrGluPheAAsnIvsAArgProGlnCysPheGluAAspThrIvsProAlaPheVal	456			
Db	1352	TACTTCTATAGAGTTTCCCAACATGCACCCGCTTTCAGAGATTTCAAGGCCACCCATATGTG	1411			
Qy	457	LyAsAlaAspHisAlaAAspGluValAArgPheValPheGlyGlyAlaPheLeuIvsGlyAAsp	476			
Db	1412	AAGGCTACACATGGTGAATGAAATTTTCTGTCTTGGCTACACAGTTC-----GGTAC	1465			
Qy	477	IleValMetPheGluGlyAlaThrGluGluIvsIleuLeuSerAArgIvsMetMetIvs	496			
Db	1466	ATAAAACTT-----CCTTACACATGAGGAAGGAGCACTTAAGCAGAAAGATATATAGAG	1519			
Qy	497	TyrTrpAlaThrPheAlaAArgThrGlyAAsnProAsnGlyAAsnAAspLeuSerLeuTrpPro	516			
Db	1520	TACTGGGCAACTTGTGCACGGCACGGGAACCCCAACAGTGAAGGCTTCACTTACGGGCT	1579			
Qy	517	AlaIvsAAsnLeuThrGluGlnIvsIleuAAspLeuAAsnMetSerLeuGlyAlaArg	536			
Db	1580	GTGATGAGCACTGATGACAGTACCTGACGCTGACATCCAGCCTTCTGTGGCTGAGCC	1639			
Qy	537	LeuIvsGluProAArgValAAspPheTrpHisSerThrIlePro-----LeuIleu	553			
Db	1640	CTGAAGGCAAGAGGCTGACGTTCTGCAGCAAAAGCTGTGCCCAAGATATCCAGAGCTA	1699			
Qy	554	SerAlaSerAAspMetLeuHisSerPro-----	562			
Db	1700	AAGGGGTCTCAGAGAGAGCAAAAGACTTGTAGTCCCTGTGTGAGAAAGAGCATGTGGG	1759			
Qy	563	-----LeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhe	575			
Db	1760	CTTCACTGTCTGAGGTTTGGAAAGCTTTTGAACCTTACCATTCACACAAACCATTC	1813			

ACCESSION	PRECURSOR (BC 3.1.1.1) (ALI-ESTERASE) (B-ESTERASE) (MONOBUTRYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYL-BUTYRASE) (Nesocercetus avaricus), full insert sequence.
VERSION	AK033563
KEYWORDS	AK033563.1 GI:26329244
SOURCE	HTc, CAP trapper.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Mamariya, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumalacostrata; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishire, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE	Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2661)
AUTHORS	Adachi, S., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shikata, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yaeunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submmission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to

Db 1438 ATMAAATT-----CCTTACACTGAGAGAGACCACTAAGCAGAGATGATGAG 1491
 QY 497 TTTTTPAAThPhaAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTyrPro 516
 Db 1492 TACCGGACCACTTTGACAGCGACCGGACCCCAACAGTAGAGGCTTACCTTACCTG 1551
 QY 517 AATATAsnLeuThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArg 536
 Db 1552 GTGATGACCACTGATGAGACGATACCTGACGCTGACATCAGACCTTCTGTGGTGGAGCC 1611
 QY 537 LeuLysGluProArgValAspPheTyrPheSerThrIlePro-----LeuIleLeu 553
 Db 1612 CTGAAGGCCAGAGAGCTGACGTTCTGACCAAGACTCTGCCAGAGATCCAGAGACTA 1671
 QY 554 SerAlaSerAspMetLeuHisSerPro----- 562
 Db 1672 AAGGGGTCTCAGAGAGCACAAGAGTTGATGCTGCTGTGAGAAAGAGCATGTGGG 1731
 QY 563 -----LeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhe 575
 Db 1732 CTTCATGCTCTGAGGTTTGGAAAGCTTTTGACCTTACCTTACACACACATTC 1785
 RESULT 3
 AY414461 1680 bp DNA linear GSS 17-DEC-2003
 LOCUS Mus musculus CES2 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY414461 GI:39770423
 VERSION AY414461.1
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1680)
 Authors Clark,A.G., Gnanoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanendau,D.M., Civeleo,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Title Interfing nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Journal Science 302 (5652), 1960-1963 (2003)
 PUBLISHED 14671302
 REFERENCE 2 (bases 1 to 1680)
 Authors Clark,A.G., Gnanoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanendau,D.M., Civeleo,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Title Direct Submission
 Journal Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source 1..1680
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1680
 /gene="CES2"
 /locus_tag="HCM5231"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,136-124 Length: 1680
 Score: 1264.00 Matches: 264
 Percent Similarity: 61.96% Conservative: 83
 Best Local Similarity: 47.14% Mismatches: 181
 Query Match: 41.05% Indels: 32
 Db: 9 Gaps: 10
 US-10-023-515-2 (1-581) x AY414461 (1-1680)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTyrPhePheLeuIleLeuGlnPro 20
 Db 16 CTTCCTGATAGGCTGAGTCTGGGCC-----TGGGGCTCTGGTT----- 57
 QY 21 LeuLeuGlyHisArgGlnTyrPheLysThrGlyProSerAlaGlnGlyProGlnArgAsn 40
 Db 58 CTCCCTCTCCATGTAGAG-----GGCCAGAGACTCAACAGTCCCTCAGAGACC 105
 QY 41 ThrArgLeuGlyTyrPheGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
 Db 106 ACACACACAGAGCCAGATCTGGCAGCCTTATCCATATAGAGACCTTGATGGGTGTC 165
 QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
 Db 166 CACTCTTCTCTGGAGATCCCTTTGCCAAGCTCTCTGATAGACCACTGGGCTTTGCACC 225
 QY 81 ProGlnProAlaSerProTyrAspAsnLeuArgGlnAlaThrSerTyrProAsnLeuGly 100
 Db 226 CCTGAGCCACCTGAACCAATGGGGTGTGTGAGAGATGGGACATCCATCCGGCCATGTGT 285
 QY 101 LeuGlnAsnSerGluTyrPheLeuLeuAspGln---HisMetLeuLysValHisTyrPro 119
 Db 286 CTGCAGACATCATCTGCATGATGATATGACAGCTTTTAAAGCTGTAGAGCTGACCTTGCTT 345
 QY 120 LysPheGlyValSerGlyAspCysLeuTyrLeuAsnIleTyrAlaProAlaHisAlaAsp 139
 Db 346 CCTTCCCATGTCTGAGAGAGCTGCTTATACCTTAACCTTAAGCACCAGATCATGCCAT 405
 QY 140 ThrGlySerLysLeuProValLeuValTyrPheProGlyGlyAlaPheLysThrGlySer 159
 Db 406 GAGGGTTCACACTCTGCTGATGTGTGATCCACAGGTGGTGGACTGGTTCAGAGCATAG 465
 QY 160 AlaSerIlePheAspGlySerAlaLeuAlaIleTyrGlnAspValLeuValValVal 179
 Db 466 GCTTCATGATATGATGATATCCGTCTGCAGACCTGAGATGTGTGTGTCTATTC 525
 QY 180 GlnTyrArgLeuGlyIlePheGlyPhePheThrTyrTyrAspGlnHisAlaProGlyAsn 199
 Db 526 CAGTACCCGCTGGGTGTCTGTGGTTCTTCAGACATGGGGACACACATGCCAGAGGCAT 585
 QY 200 TTPAlaPheLysAspGlnValAlaAlaLeuSerTyrValGlnLysAsnIleGluPhePhe 219
 Db 586 TGGGATACCTGAGACCAAGTGGCTCCCTACCTGGTCCAGCAAGATATCCCCCACTT 645
 QY 220 GlyGlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSer 239
 Db 646 GTGGCAACCGTGACCGGAGTTACATTTTGGCGAGTCTGCAGGTGGCACAAGTGTGCTT 705
 QY 240 SerLeuIleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGlnSerGly 259
 Db 706 TCACCTGTGTGTCTCCCATGTCCCAAGGGCTTTCCATGTGGCCATCATAGAGAGTGG 765
 QY 260 ValAlaIleIleProTyrLeuGlnAlaHisAspTyrGlnLysSerGlnAspLeuGlnVal 279
 Db 766 GTGGCCCTGTGCTTACTTATCTCAACACCTCGAGAGTGTCTC-----ACTAG 819
 QY 280 ValAlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArg 299
 Db 820 GTGGCCAGCTCTCTGATGTGAGGCTGAGCTGAGAGGCCCTGTGCGCTGTGTGAGA 879
 QY 300 ThrLysProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 319
 Db 880 GGCAGAGTGAAGCAGAGATTTGTGCTTAAACAAAGCTTCACATATATCCCTGTGTG 939
 QY 320 ValAspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLys 339
 Db 940 GTGATGAGAGATTCCTACCCCAAGCATCTCAGAGAGCTGTGGCTCTGTGATTTTCAAC 999
 QY 340 AlaIleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLys 359
 Db 1000 CCTGTCCCAAGCATCATTTGTTTCAACATATGATGTGTGATGTGTCCCAAAAGC 1059
 QY 360 GluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsn 379

```

Db      1060 ATGAGC-----ATGATTGACTCCAGAGGGGATGAGACAGACAAATGGTCAGACT 1110
Oy      380 ILeuHis-----1LeProGlnTyrIleuHisLeuValAla 392
Db      1111 ATCTGGAGAGAGAGGAGGACACCCAGATGATGCTGCTGCTGAGTGAATGACCTGTAATG 1170
Oy      393 AaGlnTyrPheHisAerLysHisSerLeuThrGlnIleArgHisSerLeuAaAplyLeu 412
Db      1171 GAAGAGTACATGAGGGGACACTGAGGATGCCAGCCCTCCAAATACAGTTTACAGAGATG 1230
Oy      413 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisAspAla 432
Db      1231 ATGGAGATTTCATGTTTGATGCCCTGCACTCCAAAGTACATTTT---CAGAGTTCC 1287
Oy      433 GlyAlaProValTyrPheTyrGluPheArgHisArgProGlnIysPheGluAplyThrLys 452
Db      1288 CAGCCCCCTCTACTTCTTGAATTCACATCAATCCAGCTTCTCTCAAGAGATGCAG 1347
Oy      453 ProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGlyAlaPhe 472
Db      1348 CCACCCACGATGAAACCTGACATGAGATGAGGTTCTTTGTTTGGTTCTTCTTTC 1407
Oy      473 -----LeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeu 490
Db      1408 TGGGGCATATAACTTACTTC-----ACTGAGAGAGAGAGAGCTGCTG 1449
Oy      491 SerArgLysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
Db      1450 AACAGAGAGATGATGAAAGTACAGGCACTTGCACGACATGGAAACCCCAACAGTGA 1509
Oy      511 AapLeuSerLeuTyrProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsn 530
Db      1510 GGCTACCTACTGCTGTGTGATGAGCCATGATGAGAGTACCTGACAGTGCATCCAG 1569
Oy      531 MetSerLeuGlyGlyArgLeuLysGluProArgValaAapPheThrSerThrIlePro 550
Db      1570 CTTTCTGGGTGTCGAGCCTTGAAAGCCAGAAAGGCTGCACTTGGACCAAGACTCTGCC 1629

RESULT 4
AK078953      1958 bp      mRNA      linear      HTC 03-APR-2004
LOCUS      AK078953
DEFINITION      Mus musculus adult male cecum cDNA, RIKEN full-length enriched
                library, clone:9130231C15 product:similar to LIVER CARBOXYLESTERASE
                PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus], full insert
                sequence.
ACCESSION      AK078953
VERSION      AK078953.1 GI:26347654
KEYWORDS      HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 Carninci, P. and Hayashizaki, Y.
                High-efficiency full-length cDNA cloning
                Meth. Enzymol. 303, 19-44 (1999)
JOURNAL      9979253
MEDLINE      10349636
PUBMED
REFERENCE      2
AUTHORS      Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
                Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
                Normalization and subtraction of cap-trapper-selected cDNAs to
                prepare full-length cDNA libraries for rapid discovery of new genes
                Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL      20499374
MEDLINE      11042159
PUBMED
REFERENCE      3
AUTHORS      Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
                Kono, H., Akiyama, T., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M.,
                Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A.,
                Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kasahwagi, K.,
                Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,

```

```

TITLE      Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
                Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
                RIKEN integrated sequence analysis (RISA) system--384-format
                sequencing pipeline with 384 multipillarary sequencer
JOURNAL      20530913
MEDLINE      11076861
PUBMED
REFERENCE      4
AUTHORS      The RIKEN Genome Exploration Research Group Phase II Team and the
                FANTOM Consortium.
                Functional annotation of a full-length mouse cDNA collection
                Nature 409, 685-690 (2001)
JOURNAL      5
REFERENCE      5
AUTHORS      The FANTOM Consortium and the RIKEN Genome Exploration Research
                Group Phase I & II Team.
                Analysis of the mouse transcriptome based on functional annotation
                of 60,770 full-length cDNAs
                Nature 420, 563-573 (2002)
JOURNAL      6
MEDLINE      12051002
PUBMED
REFERENCE      6
AUTHORS      Adachi, J., Aizawa, K., Akiyama, T., Arakawa, T., Bono, H., Carninci, P.,
                Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
                Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
                Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
                Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
                Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
                Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Ohata, N.,
                Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
                Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
                Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
                Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
                Muramatsu, M. and Hayashizaki, Y.
                Direct Submission
                Submitted (16-APR-2002) Yoshinide Hayashizaki, The Institute of
                Physical and Chemical Research (RIKEN), Laboratory for Genome
                Exploration and Chemical Research (RIKEN), Laboratory for Genome
                Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
                RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
                Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
                URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
                Fax: 81-45-503-9216)
                cDNA library was prepared and sequenced in Mouse Genome
                Encyclopedia Project of Genome Exploration Research Group in Riken
                Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                Division of Experimental Animal Research in Riken contributed to
                prepare mouse tissues.
                Please visit our web site for further details.
                URL: http://genome.gsc.riken.jp/
                Location/Qualifiers
                1. 1958
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="C57BL/6J"
                /db_xref="PANTOM:DB:9130231C15"
                /db_xref="PANTOM:10090"
                /clone="9130231C15"
                /sex="male"
                /tissue_type="cecum"
                /clone_id="RIKEN full-length enriched mouse cDNA library"
                /dev_stage="adult"
                65. 1742
                /note="unnamed protein product; putative
                stimilar to LIVER CARBOXYLESTERASE PRECURSOR (EC 3.1.1.1)
                [Mesocricetus auratus] (SWISSPROT|064419, evidence: FASTV,
                78.4kD, 100%length, match=1674)"
                /codon_start=1
                /protein_id="BAC37476.1"
                /db_xref="GI:26347655"
                /translation="MPILRLPGMLCVVACGILLILQHVHGQDSAPRIINTRGVGRGS
                FVHVVDKSGVHAFGIPFAPKPVGLRFPAPPEPSPGVRDGTSPAMCLOPDINW
                LEDAKENDILIPRISMSBDCILNIIYTPHAGDSNIPVWVTHIGGLVAGSASMDNY
                SKLATETIVVAIOYRLGVTGPRSDOHRGKGVGLDYVAALRQKAIATFGNR
                DRVITFGSAGCTVSIIISPMKGLFHGIHQSGVALPDLISDTSEVVAKTVANL
                SGCAITSEALIHCLRAKSKOETLIATINQVFKTIPAVVDGEPLPKHPELITLSDPHFV

```

CDS

FEATURES

Source

COMMENT

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

PSIIQVNTDSCGMVPMFMGLDHIINKITRETLPAVLKNTAARMMLPECSHLLIVEEY
MGDDTDEPFLQAFQEMFLGDPEFVTPALQVAFQSOAPVRYEEOHLSFKHYRPS
HYKADHGDVAFVPSGYLMDMLDLTEREELIKRMMCYMANPANGNPSGLPSWP
VLDDHBOYLQDLPQAVGRALKAKRLQFRTKTLPOKIGELKSOQKNAEL"

polyA_signal

1937..1942

polyA_site

/note="putative"

ORIGIN

Alignment Scores:

Pred. No.:	1,516-121	Length:	1958
Score:	1236.00	Matches:	261
Percent Similarity:	58.97%	Conservative:	81
Best Local Similarity:	45.00%	Mismatches:	187
Query Match:	40.14%	Indels:	51
DB:	3	Gaps:	8

US-10-023-515-2 (1-581) x AK078953 (1-1958)

QY 31 GYProSerAlaGluGlyProGlnIaArgAnThrArgLeuGlyTTPrlGlnGlyGln 50
DB 141 GGGCCAGAGCTCAGCCAGCCCATCAGAACCCACAGGGGTCAAGTGAAGGCGCTTT 200
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 201 GTCCACGTGAAGACACTAAATCTGTGTCATGCTTCTGGAAATCCCTTTGCCAAG 260
QY 71 ProProLeuGlySerLeuArgPheThrAnProGlnProAlaSerProTTPraAsnLeu 90
DB 261 CCTCTGTAGAGCTACTGGCTTTGACCTCTGAGAACCTGAGACATGAGATGGTGTGG 320
QY 91 ArgGluAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGluTTPleuLeuAsp 110
DB 321 AGAGATGGAGCTCAGCCGCGCATGTGTCTGCAAGCTGATATGATGATTAAGAGAT 380
QY 111 GlnHisMetLeuValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeu 130
DB 381 GCGAAGAAATGATATCTGATCTGCTCCCATCTCATGTCCGAGAGCTGCTGTATCTC 440
QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerTyrLeuProValLeuValTTPhe 150
DB 441 AACATCTACACCAACCAACGCGCAGAGGGCTCTAACCTGCTGTGATGTATGATC 500
QY 151 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
DB 501 CACGGTGTGAGCTGTGTGATGATCAGCTTCATGATGATGATGATGATGATGATGATG 560
QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
DB 561 ACTGAGAAATGATATGATGCTCCCATCCAGATCCGCTTGGGGTCTTGGCTTCTTACGC 620
QY 191 ThrTTPraSerGlnHisAlaProGlyAsnTTPraLapheLysAspGlnValAlaLeuSer 210
DB 621 ACTGAGAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 680
QY 211 TTPValGlnLysAsnIleGluPheGlyGlyAspProSerSerValThrIlePheGly 230
DB 681 TGGGTGCGAAGAAATGATGCTTATTTGGAGCAACCGTGAAGGGTCTCTATATTTTGA 740
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaValGlyLeu 250
DB 741 GTCTAGCGAGGAGTACAGATGTGTCTTCTCATATCTTCTCCCAATGTCCAAAGACCTC 800
QY 251 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyrIleGluAlaHisAsp 270
DB 801 TTTTATGAGGCACTATATGAGAGTGAAGTGGCGCTGCTGCTGCACTATCTCT----- 854
QY 271 TyrGluLysSerGluAspLeu---GlnValValAlaHisPheCysGlyAsnAsnAlaSer 289
DB 855 ---GACACTTCTGAGGTGTCTTACAAAGACATGAGCACTATCTGTTGTGAAACCA 911
QY 290 AspSerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuThrLeu 309

DB 912 GACTCAGAACCCCTGATCATGCTGCTGAGAGCAAGAGTAAACAAGATCTGCTGATTT 971
QY 310 SerGlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnLysPro 329
DB 972 AACCCAGTCTTCCAAATGATGCTGCTGTGGATGGAGAGTTCTTACCCAAAGATCTT 1031
QY 330 LeuAspLeuLeuSerGlnLysAlaPheLysValAlaIleProSerIleIleGlyValAsnAsn 349
DB 1032 CAGAGCTGCTGACCTCATGATGATTTTCAACCTGCTGCCAGATCATTTGTCTCAACACA 1091
QY 350 HisGluCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsn 369
DB 1092 GATGATATGGCTGGAGGATGCCCATG-----TTCATGGGCTTGTAC 1133
QY 370 LysSerLeuAlaLeuHisLeuIleGlnAsnIle----- 380
DB 1134 -----CATATCATAAAGAACATTAACACAGAGACCTTGCAGCTTGTG 1178
QY 381 -----LeuHisIleProProGlnTyrLeuHisLeuValAlaAsnGlu 394
DB 1179 AAGAACACAGACACCGAATGATGCTGCTCCAGAAATGATGATGATGATGATGATG 1238
QY 395 TyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAspLeuGly 414
DB 1239 TATATGGGGGACACTGAGAGACCCAGAGACCTGCAAGACAGTTCAAGAGATGCTGGGG 1298
QY 415 AspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAla 434
DB 1299 GACTTCATGTTGTGATGCTGCTCTCCAAAGTGGACATTTCCAAAGCT---TCCAGAGCT 1355
QY 435 ProValTyrPheTyrGluPheAlaHisArgProGlnCysPheGluAspThrLysProAla 454
DB 1356 CCTGTACTCTTATGAGTTCAGAGCTGTCAGACTGTCAGCTTCAACAGATGTCAGGCATCA 1415
QY 455 PheValValAlaAspHisAlaAspGluValArgPheValPheGlyGlyAlaPheLeuLys 474
DB 1416 CATGGAAGGACAGACCAAGATGATGAGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTG 1475
QY 475 GlyAspIleValMetPheGluGlyAlaThrGluGluGluLysValLeuLeuSerArgLysMet 494
DB 1476 ATGAATTTGGACCTC-----ACTGAGGAGGAGAGCTGCTTAAAGGATGATG 1523
QY 495 MetLysTyrTTPraLalaThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeu 514
DB 1524 ATGAAGTACTGGGCAATTTGCAAGAAATGGAACCCCAACAGTGAAGGCTTACCTTCC 1583
QY 515 TTPProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGly 534
DB 1584 TGGCTGTGCTTGGACCAAGATGATGAGCAATCTGCAAGCTGAGACACCAAGCTGCTGGGT 1643
QY 535 GlnArgLeuLysGluProArgValAspPheTTPraSerThrIleProLeuIleLeuSer 554
DB 1644 CGAGCCCTGAAGGCGCAAGAGGCTGAGTTCGACCAAGACTGTCCTCCCAAGAAATCCAG 1703
QY 555 -AlaSerAspMetLeuHisSerProLeuSerSerLeuThr----- 567
DB 1704 GAGCTCAAGGTTCTCAGAGACAAGCATCAGAGCTGTGAGCTGCTGTCTCAGGAAAGTGT 1763
QY 568 -----PheLeuSerLeuLeuGlnProPhePhePhe 578
DB 1764 TGTGAGATTTAAAGACAGATGCTGATTTCTGAGGTTACAAAGTCAATTTATTCATTT 1821

RESULT 5

AY144459

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1680 bp DNA linear GSS 17-DEC-2003
Homo sapiens CES2 gene, VIRIUTL TRANSCRIPT, partial sequence,
genomic survey sequence.

AY144459
AY144459.1 GI:39770421
GSS.
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1680)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A., Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL PUBMED 14673302
2 (bases 1 to 1680)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A., Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source 1. 1680
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1680
/gene="CES2"
/locus_tag="HCM5231"
ORIGIN
Alignment Scores:
Pred. No.: 5,25e-121 Length: 1680
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 39.95% Indels: 24
DB: 9 Gaps: 8
US-10-023-515-2 (1-581) x AY14459 (1-1680)
QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPileGlnGlyLeuGln 50
DB 76 GGGCCAGACTCAGCCAGTCCATCCGACACACACAGGGGACAGCTCGGGAGTCTT 135
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 136 GTCCATGTGAAGGGGCGCAATGCCGGGGTCCAAACCTTCCTGGGAATTCATTGCCAAG 195
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu 90
DB 196 CCACCTCTAGGTCGCGTGCAGATTGCACCCCTGAGCCCTGAAATCTTGGAGTGGTGTG 255
QY 91 ArgGlnAlaThrSerTyrProAsnLeuGlyLeuGlnAsn-----SerGln 105
DB 256 AGGATGGAACACCAATCCGCGCATGTGTACAGACCTCACCGACAGTGGAGTCAAG 315
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuValHisTyrProLysPheGlyValSerGln 125
DB 316 TTTTCT-----AGCCAGTTCACATGACCTTCCTTCGACTCCAGTCAATGTGAG 363
QY 126 AspCysLeuTyrLeuAsnIleTyrAlaProAlaHisAlaAspThrArgLysSerLysLeuPro 145
DB 364 GACTGCTGTACCTCAGACATCTACACCCCGGCGCATAGCCATGAGAGCTTAACTGCGC 423
QY 146 ValLeuValTyrPheProGlyGlyValaPheLysThrGlySerAlaSerIlePheAspGly 165
DB 424 GTATATGTGTGGATCCACCGGTGTGCGCTGTTTGGCATGCTTCCTTGATGATGCT 483
QY 166 SerAlaLeuAlaIleTyrGlnAspValLeuValValValaGlnTyrArgLeuGlyIle 185
DB 484 TCCATGTGCTGCTCTTGAGAAAGTGTGTGTGTATCATCAAGTACCGCTCGGGTGTTC 543

QY 186 PheGlyPhePheThrThrTyrAspGlnHisAlaProGlyValAsnTyrAlaPheLysAspGln 205
DB 544 CTGGGCTTTCAGACTGAGACAGACAGACGACACCGGCACTGGGGCTACCTGGACCA 603
QY 206 ValAlaAlaLeuSerTyrValGlnLysAsnIleGluPheGlyGlyAspProSerSer 225
DB 604 GTGGTGCATCAGCTGGGTCCAGAGATATCCGCCACTTTGGAGGCCAACCCTGACCGT 663
QY 226 ValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
DB 664 GTACACATTTTGGGAGATCTCGGGGTGCACAGATGTGTCTTGCTGTGTGTCTCC 723
QY 246 MetAlaLysGlyLeuPhePheIleValaIleMetGlySerGlyValaIleIleProTyr 265
DB 724 ATATCCAAAGACTCTTCACAGAGCATCATGAGAGTGGCGTGGCTCTCCCGCC 783
QY 266 LeuGlnAlaHisAspTyrGlyLysSerGlnAspLeu-----GlnValValaHisPhe 283
DB 784 CTCAATGGC-----AGCTCAGCTGATGTATCTCCACGGTGTGGCCAACTGT 831
QY 284 CysGlyAsnAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThrLysProSer 303
DB 832 TCTGGCTGTGACCAAGTTGACTGAGGCTGTGGGTGCTCGCGGGCAAGACTAA 891
QY 304 LysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValValaAspGlyAla 323
DB 892 GAGAGATCTCTTGCAATTACMACCTTTCAGATGATCCCGGAGTGGTGGTGGGCT 951
QY 324 PhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheValaIleProSer 343
DB 952 TTCCTGCCAGGACCCCGACAGAGCTGTGCTCTGCCACTTTCAGCTGTCCCTAGC 1011
QY 344 IleIleIleValaLysAsnHisGlyCysGlyPheLeuLeuPro-----MetLys 359
DB 1012 ATGTGTGTGTCAACAATGATTCGCTGCTCATCCCAAGGATCATGAGACTTAT 1071
QY 360 GlnAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsn 379
DB 1072 GATACCCAGAAAGAAATGACAGAGAGGCTCCAGGCTGTCTTGAGAAATGTTCAC 1131
QY 380 IleLeuHisIleLeuProGlnTyrLeuHisIleValaIleAsnGluTyrPheHisAspLys 399
DB 1132 CTGCTGATGTGCTCTCTCATTTGTGTGACCTGTGAGGAGAGATACATTGGGACAT 1191
QY 400 HisSerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
DB 1192 GGGGATCCCAAGCCTCTCAAGGAGATTCCAGAGATGATGGCGGACTCATGTTTGTG 1251
QY 420 ValProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
DB 1252 ATCCCTGACCTCAAGTACATCTTT---CAGTGTCCCGGGCCCTGTGTACTTCTAC 1308
QY 440 GluPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAsp 459
DB 1309 GAGTTCACACATCAGCCGCTGCTCAAGAACATCAGGCCACCGCACATGAAGCAGAGC 1368
QY 460 HisAlaAspGluValaArgPheValaPheGlyGlyAlaPheLeuLysGlyAspIleValMet 479
DB 1369 CATGTGATGAGCTCTTTTGTTC---AGAAATTTCTTTGGGGGCAACTCATCAATAA 1425
QY 480 PheGluGlyValaThrGlnGlnGluLysLeuLeuSerArgLysMetMetLysTyrTyrAla 499
DB 1426 TTC-----ACTGAGGAAGAGAGAGCTTAAGCAGAAAGATGATGATTAATCTGGGCC 1476
QY 500 ThrPheAlaArgThrGlyAsnProAsnGluLysAspLeuSerLeuTyrProAlaTyrAsn 519
DB 1477 AACTTGGAGAAATGGAAACCCCAATGGCCAGAGGTCTGCACACTGGCCCTCTTGAC 1536
QY 530 LeuThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlu 539
DB 1537 CAGAGGAGCAATACCTGCACTGAACCTTACAGCTGTGGGTGGCCGGGCTCTGAAGGCC 1596
QY 540 ProArgValaAspPheThrPheThrIlePro 550

Db 1597 1597 CAGAGCTTCAGTTCTCGAAGAGGCGCTGCC 1629

||||:| ||||| :|||

RESULT 6
AK040349

LOCUS
AK040349

DEFINITION
Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone: A330088B12 product: similar to CARBOXYESTERASE PRPCUNSOR (EC 3.1.1.1) (AL-ESTERASE) (B-ESTERASE) (MONOBUITYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full insert sequence.

ACCESSION
AK040349

VERSION
AK040349.1 GI:26087790

KEYWORDS
HTC; CAP trapper.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y

TITLE
High-efficiency full-length cDNA cloning

JOURNAL
Meth. Enzymol. 303, 19-44 (1999)

MEDLINE
99279253

PUBMED
10349636

REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE
20499374

PUBMED
11042159

REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsubara, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer

JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE
20530913

PUBMED
11076861

REFERENCE
AUTHORS
4

TITLE
The RIKEN Genome Exploration Research Group Phase II Team and the Functional Consortium.

JOURNAL
Functional annotation of a full-length mouse cDNA collection

REFERENCE
AUTHORS
Nature 409, 685-690 (2001)

TITLE
5

JOURNAL
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

REFERENCE
AUTHORS
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420, 563-573 (2002)

REFERENCE
AUTHORS
6 (bases 1 to 4927)

JOURNAL
Adachi, J., Aizawa, K., Akiyama, T., Arahawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hoti, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsumura, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takeuchi-Ahara, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

REFERENCE
AUTHORS
Direct Submission

JOURNAL
Submitted (16-JUL-2001) Yoshitake Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

COMMENT

RIKEN Yokohama Institute; 1-7-22 Suifu-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail:genome-research@riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

Location/Qualifiers

FEATURES

Source

1..4927

organism="Mus musculus"

mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM.DB:A430088E12"

/db_xref="taxon:10090"

/clone="A430088E12"

/tissue_type="thymus"

/clone_lib="RIKEN full-length enriched mouse cDNA library /dev stage=0 day neonate"

25..1711

/note="putative

similar to CARBOXYL ESTERASE PRECURSOR (BC 3.1.1.1) (AL-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLESTERASE) (Mecroicetetus auratus) (SPTR|035533, evidence: FASIV, 70.7%id, 99.2%length, match=1672)"

4912..4917

/note="putative"

4927

polyA_site

polyA_signal

polyA_site

/note="putative"

ALIGNMENT SCORES:

Prod. No.:

Score: 1.06e-116 Length: 4927

Percent Similarity: 61.98% Matches: 248

Best Local Similarity: 47.15% Mismatches: 78

Query Match: 36.88% Indels: 17

DB: 3 Gaps: 6

US-10-023-515-2 (1-581) x AK040349 (1-4927)

QY 33 SerLaGlugLyProGInARgAsnThRArgLeuGlyTrPIIeGInGlyLysGInValThr 52

Db 112 TCAGAGGCCAGCCCCCATCAGAAACACATVCTGACAGGTCGAGGCGAAAGTTGTCCAC 171

QY 53 ValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIaIaProPro 72

Db 172 TTGACAGACATCAAGAAGCGAGGCCATTAATCTTCTGGGAGATCCCTTTGCCAAGCTCT 231

QY 73 LeuGlySerLeuAArgPheThrAanProGInProAlaSerProTrPaaPaaLeuAArgG 92

Db 232 GTAGAGCAACACTGCCTTTGACACCTCTCGAGGCCCTTGAAGCATGAGTGGTGAGAGA 291

QY 92 uAlaThrSerTryrProAsnLeuCySleuGInAAsSerGluTrPheuLeuLeuAap--G 111

Db 292 TGGGACCTTCAGAGCGCCGATGTGTCTACAAATGATGATVATGGAATTTAGAGGGTCT 351

QY 111 nHiMetLeuLyValIaHTryrProLyPheGlyValSerGluAapCySleuTrLeuAs 131

Db 352 GAAATATATAAAGATGATCCCTGCTCTCTCATGCTGAGAGCTCCGTATATCTCA 411

QY 131 nIleTrAlaProAlaHiSaIaAapTrngLySerLySleuProValLeuValTrpPhePr 151

Db 412 CATCTTACACACCAAGCCAGCGCCAGAGGGGCTTAACCTGCTGTGATGTGTGGCTCCA 471

QY 151 oGlyGlyValaPheLySrHngLySerAlaSerIlePheAapGlySerAlaLeuAlaIaTr 171

Db 472 TGTGTGTGAGACTGGTTCAGAGCATGAGCTTTCATATATGATGATTCCTGTGCGAGCCAC 531

Score: 1191.00 Matches: 242
Percent Similarity: 62.15% Conservative: 88
Best Local Similarity: 45.57% Mismatches: 161
Query Match: 38.68% Indels: 40
DB: 3 Gaps: 8
US-10-023-515-2 (1-581) x HSM803013 (1-2888)

QY 31 GYProserAlaGluGlyProGlnArgAntThrArgLeuGlyTyrPrlLeuGlnIlySerGln 50
DB 1221 GCCCAGAGCTCCAGCTCCATCCGACACACACACGCGGCAAGTGTGCTGGAGTCTT 1280
QY 51 ValThrValLeuGlySerProValProValAenValPheLeuGlyValPhePheAla 70
DB 1281 GTCCATGTGAAGGGGCCCATCCGGGCTCAACCTTCTGGAAATTCATTGGCAAG 1340
QY 71 ProPheLeuGlySerLeuArgPheThrAnProGlnProAlaSerProTrrPheAenLeu 90
DB 1341 CCACCTTAGGTCGCGCTGGATTGTGACCCCTGAGCCCTGAAATCTTGAGTGGTGTG 1400
QY 91 ArgGluAlaThrSerTyrProAenLeuGlyLeuGlnAen-----SerGlu 105
DB 1401 AGGAGTGAACCCATCCGACATGTGTCTACAGACCTCACCGACGTGAGTCAGAG 1460
QY 106 TrpLeuLeuLeuAerGlnHisMetLeuLysValHisTyrProLysPheGlyValSerGlu 125
DB 1461 TTTCTT-----AGCCAGTTCACAGATGACCTTCCTCCGATCTCATGTCTGAG 1508
QY 126 AspGlyLeuTyrLeuAenIleTyrAlaProAlaHisAlaAerThrGlySerLysLeuPro 145
DB 1509 GACTCCCTGATCTCAGCATCTACACGCGGCCCATAGCATGAAGGCTCTAACCTGCGG 1568
QY 146 ValLeuValTrrPheProGlyGlyValAphelyThrGlySerAlaSerLlePheAerGly 165
DB 1569 GTGAGGTGTGTGATCCAGCGTGTGCTGTCTTTTGTGGCATGCTCTCTGTATATAGTGT 1628
QY 166 SerAlaLeuAlaAlaTyrGluAerValLeuValAlaValAlaGlnTyrArgLeuGlyLe 185
DB 1629 TCCATGCTGGCTGCTTGGAGAACTGGGTGTGTCTATCTCATACGATCCGCTGGGTGTC 1688
QY 186 PheGlyPhePheThrTrrPhePheGlnHisAlaProGlyAenTrrAlaAphelyAerGln 205
DB 1689 CTGGGCTTCTCAGACATGAGACACAGACACGCAACGCGCACTGGGGCTACCTGAGCCAA 1748
QY 206 ValAlaAlaLeuSerTrrValGlnLysAenIleGluPhePheGlyGlyAerProSerSer 225
DB 1749 GTGGCTGACATACGCTGGGTCCAGACATATCGCCATTGTGAGGCAACCTTACCGCT 1808
QY 226 ValTrrLlePheGlyGlySerAlaGlyAlaIleSerValSerSerLleuLleuSerPro 245
DB 1809 GTACACATTTTGGGAGTCTGGGGTGGCACGAGTGTCTTCGCTGTGTGTGTCTCCCC 1868
QY 246 MetAlaLysGlyLeuPheHisLysValAlaIleMetGlySerGlyValAlaIleLeuProTyr 265
DB 1869 ATATCCCAAGACCTTCCACAGGACCATCATGAGAGTGGGTGGCCCTCTGCGCGGC 1928
QY 266 LeuGluAlaHisAerTrrGlyLysSerGluAerLeu-----GlnValAlaHisAerPhe 283
DB 1929 CTCATATGCC-----AGCTCAGCTGATGTCTATCTCCACGCTGTGTGCGCAACCTG 1976
QY 284 CysGlyAenAenAlaSerAerSerGlnAlaLeuLeuAerGlyLeuAerGlnArgProSer 303
DB 1977 TCTGCGCTGTACAAATTGACTGTAGGCTGTGGGCTGTGGGCTGTGGGCGCAAGATAA 2036
QY 304 LysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValAlaAerGlyAla 323
DB 2037 GAGGAGATTTCTTGCAATTATCAAGCCTTCAAGATGATGCCGCGAGTGTGTGATGGGCTC 2096
QY 324 PhePheProAenGlnProLeuAerLleuLeuSerGlnLysValAphelyAerAlaLeuProSer 343
DB 2097 TTCTGCGCCAGGACCCCAAGAGCTGTGGCTGTGCGCATTTTCAGCTCTCTCCCTTAC 2156
QY 344 IleIleGlyValAerAerHisGlyCysGlyPheLeuLeuPro-----MetLys 359

DB 2157 ATGTGTGTGTCAACAACATGAATTCGGCTGGCTTCATCCCAAGATCATGAGATCTAT 2216
QY 360 GluAlaProGluIleLeuSerGlySerLysAerLleuAlaLeuHisLleuIleGlnAen 379
DB 2217 GATACCCCAAGAAATGATGACAGAGAGCTCCCAAGGCTGCTGTGACAAATGTTAAAG 2276
QY 380 IleLeuHisLleProProGlnTrrLleuHisLleuValAlaAenGlnTrrPheHisAerLys 399
DB 2277 CTGTGATGTTGCTCTTCAATTTGTGTGATCTGTGAGAGAGATTCATTGGGACAT 2336
QY 400 HisSerLleuThrGluIleArgAerSerLleuAerLleuLeuGlyAerValAerPheVal 419
DB 2337 GGGGATCCCAAGACCTCCAAAGCCGATTCACAGAGATGATGGCGATCTCATGTTGTG 2396
QY 420 ValProAlaLeuIleThrAlaArgTrrHisAerAerAlaGlyAlaProValTrrPheTyr 439
DB 2397 ATCCCTGACATCTCAAGTACATTTT---CAGTGTCTCCGGGCGCTGTGTACTTCTAC 2453
QY 440 GluPheArgHisAerProGlnLysAerPheGlnAerThrLysProAlaPheValLysAlaAer 459
DB 2454 GAGTTCACAGATCAGCCCAAGCTGGCTCAAGACATCAGGACACCGCACATGAAGGACAGC 2513
QY 460 HisAlaAerGluValAerPheValPheGlyGlyAlaPheLeuLysGlyAerLleuValMet 479
DB 2514 CAT-----GTAAATTC----- 2525
QY 480 PheGluGlyAlaThrGluGluGlyLysLeuLeuSerAerGlyLysMetLysTrrTrrAla 499
DB 2526 -----ACTGAGAAAGAGAGACAGCATGACAGATGATGATGAATGATGATGAGTGGGCT 2573
QY 500 ThrPheAlaArgThrGlyAerProAenGlyAerAerLysAerLleuTrrProAlaTrrAen 519
DB 2574 AACTTGTGAGAAATGAGAAACCCCATGACGATCTGACACACTGGCGCTGTTCGAC 2633
QY 520 LeuThrGluGlnTrrLeuGlnLysAerLysAerMetSerLleuGlyGlnAerLysGln 539
DB 2634 CAGAGAGACATACCTCAGCTGATGAACCTACAGCTGTGGGTGGCGGGCTTGAAGGCC 2693
QY 540 ProArgValAerPheTrrPheTrrSerThrLlePro 550
DB 2694 CACAGCTTCAGTCTGTGAAAGAGCGCTGCC 2726

RESULT 8
HSM806270
LOCUS HSM806270 3909 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686H0466 (from clone DKFZp686H0466).
ACCESSION BX538086
VERSION BX538086.1 GI:31874222
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3909)
AUTHORS Bloeker,H., Boecker,M., Brandt,P., Mewes,H.W., Well,B., Amid,C., Oesinger,A., Fobo,G., Han,M. and Wiemann,S.
CONSRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686H0466) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686H0466
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers

```
source
1..3909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686H0466"
/db_xref="taxon:9606"
/clone="DKFZp686H0466"
/tissue_type="endometrium carcinoma cell line"
/clone_id="686 (synonym: hlcc3). Vector pSPORT1_Sfi, host
DH10B, sites SfiIa + SfiIb"
/dev_stage="adult"
/notc="carboxylesterase 2 isoform 2"
1..3909
/gene="DKFZp686H0466"
/1002..2825
/gene="DKFZp686H0466"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD98009.1"
/db_xref="GI:31874223"
/db_xref="UniProt/Swiss-Prot:O00748"
/translations="MTAQSRSPTTPPPSPQRTPLTPCPVQTPRRKALHICWTDPG
OPLEGQVRBORRTSEPTMPLHRLARLSAVAGLLLVKGQSDSPRTTHT
GOVLSLVKVKANNAGVQTPLGIPFAKPPUGLRFAPRPPSPWSGVRDGTTHPMCL
ODLTAVSESPSPNTPSPDSMSBDCLYSTTPAHSHSGSNLPVMTWIHGALVFG
MSLIDGSMALALENVVVIQYRLVGFSTGDHATNGAYLDQVALRWQONI
AHFGNPDVITFESAGGTSVSLVSVISQLEFGALMESGVALPLGLIASADVI
STVVANLSACDQVSEALVGLGRKSEIILAIINKEFKMIPVDVGFPLPRPOLLA
SADFOVPSIVGVNNHFGMLPKVMRIYDQKEMDRBSAOLIKMTLTLMLPTSG
DLIREYIDGNDPOTLQAFQBMADSMFVTPALQVAFQCSRAPIVTFEFOHPSM
LNIRPPIHKADHVTFEERPOLSRKMTYMANPFRANGNPSGLPHMPLFDEEYOYL
RNLQPAVARALKARLPQKALPQKIQLESEPERHTEL"

ALIGNMENT SCORES:
Pred. No.: 3,05e-115 Length: 3909
Score: 1182.00 Matches: 241
Percent Similarity: 61.96% Conservative: 88
Best Local Similarity: 45.39% Mismatches: 162
Query Match: 38.39% Indels: 40
Dbs: 3 Gaps: 8

US-10-023-515-2 (1-581) x HSN806270 (1-3909)
QY 31 GTPProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTTPILeGlnGlyValSerGln 50
Db 1269 GGGCCAGAGCTCCAGCCAGTCCATCCGAGCCACACACACAGGGGCGAGTGTGAGTCTT 1328
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 70
Db 1329 GTCCATGTGAAGGGGCGCAATGCCGGGGTCCAAACCTTCTCGGAAATTCATTTGCCAAG 1388
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrPaspAsnLeu 90
Db 1389 CCACCTCTAGGTCCGCTCGCATTTGCACCCCTGAGCCCTCGAATCTGGAGTGGTGTG 1448
QY 91 ArgGlnAlaThrSerTyrProAsnLeuCyLeuGlnAsn-----SerGln 105
Db 1449 AGGGATGGAACACACCATCCGCGCATGTGTTCAGAGACCTCACCGCAGTGGAGTCCAGAG 1508
QY 106 TrpLeuLeuAsnArgGlnHisMetLeuValHisTyrProTyrPheGlyValSerGln 125
Db 1509 TTTCTT-----AGCCAGTTCAACATGACCTTCCCTTCCGATCCGATGTCGAG 1556
QY 126 AspCyLeuTyrTrpLeuAsnIleTyrAlaProAlaHisAlaAspThrArgSerLeuPro 145
Db 1557 GACGTGCTGTACCTCCAGACATCTACACGCCGCGCATGACCATGAAGGCTTAACTGCGCC 1616
QY 146 ValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGly 165
Db 1617 GTGATGTGTGGATCCACGCTGTGGCTGTGTTTGGCATGCTCTTGTATGATGTG 1676
QY 166 SerAlaLeuAlaAlaTyrGluAspValLeuValValValGlnTyrArgLeuGlyIle 185
```

```
Db 1677 TCCATGCTGCTGCTGCTTGGAGAACGNGTGTGTCATCATCCATACCCGCTGGTGTG 1736
QY 186 PheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTTPAlaPheLeuAspGln 205
Db 1737 CTGGGGTCTTCTTCAACCTGAGAACAGACCAACCGGCAACTGGGGCTGATCTGGACCA 1796
QY 206 ValAlaAlaLeuSerTrpValGlnLysAsnIleGluPhePheGlyValAspProSerSer 225
Db 1797 GTGGCTGACATACGCTGGGTCCAGACAAATATCCGCCACTTTGAGGCAACCTGACCGT 1856
QY 226 ValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
Db 1857 GTACACATTTTGGGAGCTGCGGGTGGCAAGAGTGTCTTGCTGTGTGTGTCCTCC 1916
QY 246 MetalAlaValGlyLeuPheHisLeuAlaIleMetGluSerGlyValAlaIleProTyr 265
Db 1917 ATATCCACAGACATCTTCCACGAGGCCATCATGGAGTGGCGTGTGCTGCTGCTGCGGC 1976
QY 266 LeuGlnAlaHisAspTyrGluLysSerGluAspLeu-----GlnValValAlaHisPhe 283
Db 1977 CTCAATGGC-----AGCTCAGCTGATGTCACTCCACGGTGGTGGCCAACTG 2024
QY 284 CysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuArgCysLeuArgThrLysProSer 303
Db 2025 TCTGCTGTGACCAAGTTGACTGTGAGGCCCTGTGTGGCTGCTGCGGGCAAGATGAA 2084
QY 304 LysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValValaAspGlyAla 323
Db 2085 GAGAGATCTCTTGCAATTAACAAGCTTTCAGATGATCCCGAGAGTGGTGGTGC 2144
QY 324 PhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAlaIleProSer 343
Db 2145 TTCCTGCTCCAGGACCCCGAGAGCTGTGGCTGTGCGGACATTTGACCTGTCCCTAGC 2204
QY 344 IleIleGlyValaAsnAsnHisGlyCysGlyPheLeuLeuPro-----MetLys 359
Db 2205 ATTGTGTGTCTCAACAATGAATGATTCGGTGTCTATCCCAAGTCAATGAGATCTTAT 2264
QY 360 GluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuIleGlnAsn 379
Db 2265 GATACCCAGAAAGAAATGACACAGAGAGCCCTCCAGCTGTCTTGACAAATATGTTAACG 2324
QY 380 IleLeuHisIleProProGlnTyrLeuHisLeuValAlaAsnGluTyrPheHisAspLys 399
Db 2325 CTGCTGATGTGTGCTCTCTATCATCTGTGTGACCTGTGAGGAGAGATGACTTGGGACAT 2384
QY 400 HisSerLeuThrGluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheVal 419
Db 2385 GGGGATCCCAAGACCTCCCAAGCGCAGTTCCAGAGATGATGCGGAGCTCCATGTTTGTG 2444
QY 420 ValProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
Db 2445 ATCCCTGACCTCCAAAGTACACATTTT--CAAGTGTCCCGGGCCCTGTGTATCTTAC 2501
QY 440 GluPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAsp 459
Db 2502 GAGTTCAGACATCAAGCCGAGCTGGCTCAAGAACATCAGGACACCGCACATGAGCAAGC 2561
QY 460 HisAlaAspGluValArgPheValPheGlyValaPheLeuLysGlyAspIleValMet 479
Db 2562 CAT-----GTAAATTC----- 2573
QY 480 PheGluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyrTrpAla 499
Db 2574 -----ACTGAGAAAGAGAGAGACGATTAACAGAAAGATGATGAATACCTGGGCC 2621
QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTTPProAlaTyrAsn 519
Db 2622 AACTTTGGAGAAATGAGAACCCCAATGGCGAGGATGTGCACACTGGCGCTGTGCAC 2681
QY 520 LeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlu 539
Db 2682 CAGAGAGAGCAATACCTGCGGCTGAACCTTACAGCTTGGGTGGGCGGCTGTGAAGGCC 2741
```

QY 540 ProArvaIaSpPheTrpThrSerThriIlePro 550
DB 2742 CACAGCTCCAGTCTTGAGAGAGGCGCTGCC 2774

RESULT 9
BC019926 2038 bp mRNA linear HTC 20-SEP-2002
LOCUS Mus musculus, clone IMAGE:5123923, mRNA.
DEFINITION BC019926.1 GI:18044766
ACCESSION BC019926.1 GI:18044766
VERSION HTC.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2038)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL ac: <http://image.lnl.gov>
Series: IRAX Plate: 42 Row: P Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
This clone has the following problem: frame shifted.
Location/Qualifiers
1..2038
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5123923"
/tissue_type="Liver, normal. 5 month old male mouse."
/clone_lib="NCI CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Alignment Scores:
Pred. No.: 3..7e-113 Length: 2038
Score: 1158.50 Matches: 255
Percent Similarity: 59.65% Conservative: 86
Best Local Similarity: 44.58% Mismatches: 183
Query Match: 37.63% Indels: 48
DB: 3 Gaps: 12
US-10-023-515-2 (1-581) x BC019926 (1-2038)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrp-----CysPhePheIleIleu 18
DB 32 ATACAGGCTGGG-----TCCAGTGTCTTAAGTGTGACCTGTCTGTGGCATTT 85
QY 19 GlnProLeuLeuGlyHisArgGlnTrpGlyLysThrGlyProSerAlaIleuGlyProGln 38

DB 86 GTTACACAGTC-----ACAGACCCAAAGTCATTCAGCTGAA 124
QY 39 ArgAenThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProVal 58
DB 125 GTGATACCCCGCTGGTGTGTGAGGCGCGAGGTGGTGTAGAGACACAGACCGC 184
QY 59 ProValAenValPheLeuGlyValProPheIleAlaProProLeuGlySerLeuArgPhe 78
DB 185 ATGTAAATGTCTTCTGTGGCATTCCTTTGCTCAAGCACAGTGGACCTTTCGGTTC 244
QY 79 ThrAenProGlnProAlaSerProTrpPheAenLeuArgGlnAlaThrSerTrpProAen 98
DB 245 TCAGCTCCATCCCAACACAGCCCTGGGAAGGTGTAGAGATGCCAGCATCAATCCCCA 304
QY 99 LeuCyLeuGlnAenSerGlyTrpLeuLeuLeuAenGlnHisMetLeuLysValHisTrp 118
DB 305 ATGTGCTTCAGAGATGTAGAAATGATCAACAGCATTTGGCTCAATGAAAGATT 364
QY 119 ProLysPheGlyValSerGlyLeuProValLeuValTrpPheProGlyValAlaPheLysThrGly 138
DB 365 AAAATCTTCCCATTTCTGAGACTGTCTGACCTCAACATCTACAGCCCACTGAGATC 424
QY 139 AspThrGlySerLysLeuProValLeuValTrpPheProGlyValAlaPheLysThrGly 158
DB 425 ACTGACAGGGGACAAAGGCGGTGATGATGATTCACAGAGGCTCTGTGGTGGC 484
QY 159 SerAlaSerIlePheAspGlySerAlaLeuAlaIleTrpGlyLeuValValVal 178
DB 485 TCTTCACATCTCAGAGATGTAGTACACATCGCGGGGTAGTGTATTTGCTACT 544
QY 179 ValGlnTrpArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHisAlaProGly 198
DB 545 GTTCAGATTCGCTTGGGATCTTGCTTCTTCACAGCTGGAGACAGCATGCCAGGC 604
QY 199 AenTrpAlaPheLysAspGlnValAlaLeuSerTrpValGlnLysAenIleGluPhe 218
DB 605 AACAGGGGATCTGTGATGTGTGTGCTGTCTTCCTGGGTGCAGGGAACTAGGCCCC 664
QY 219 PheGlyGlyAspProSerSerValThrIlePheGlyLysSerAlaGlyValIleSerVal 238
DB 665 TTTGGGGGGGTATCCCACTGTGTCAATCTTCGTGATCTCTGTGGAGGATGTTGT 724
QY 238 LysSerLeuIleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGluSe 258
DB 725 CTCAATCCGTTCTGTCTCCATATCTGTGGGCTTTCACAGAGCCATATCCCAAG 784
QY 258 TGIYValAla-----IleIleProTrpLeuGlnAlaHisAspTrp 271
DB 765 TCGGATGTGACAGACATTATGATGAGACATGAACCATGCGCTGAGACTCAGACTT 844
QY 271 TGIYLysSerGlyAspLeuGlnValAlaHisPheCyGlyLeuAenAlaSerAspSe 291
DB 845 TCCCAATCT-----GTGGCC-----TGTGCTCCCAAGC 880
QY 291 TGIYAlaLeuLeuArgCyLeuArgThrLysProSerLysGlyLeuLeuThrLeuSerG 311
DB 881 TTAG--CTGTGTCAGTGTGTGTCAGAGAGAGAAAGACCTT-----ATCAAGCA 931
QY 311 nLysThrLysSerPheThrArgValAlaAspGlyAlaPhePheProAenGlnProLeuAs 331
DB 932 GAAAGACGTAACATTCTCTACATAGTCAATGACTCTTCTCCCAAGAGCCAGAGAA 991
QY 331 PheLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAenAsnHisG 351
DB 992 GCTCTTAGACAGACCAATTCCTCCACGTGCTTACCTTTGGAGATGCACCAACATGTA 1051
QY 351 uCyGlyPheLeuLeuProMetLys-----GluAl 361
DB 1052 ATTGAGCTGGCTTCACTCAATCTTGAATATCTGTGATTAAGTTGAACATCTGAGCCG 1111
QY 361 aProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisIleuIleGlnAenIle 381

Db 133 CTGTAAGTCACAGACGCTTAGGT-----GCTTTGACACCAAG 171
 Qy 36 G1YProGlnArgAsnThrArgLeuGlyTyrIleGlnGlyValThrValLeuGly 55
 Db 172 GAACCTCTCTCATATCAACACGAGATCTCCAGGAAGAACAGATGATG----- 225
 Qy 56 SerProValProValAsnValPheLeuGlyValProPheAlaAlaProProLeuGlySer 75
 Db 226 GGAGACACACCCATCAAGTCTTCCGAGGAGATCCCTTCTTAACCTCCGGTGGGCAAC 285
 Qy 76 LeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeuArgLysAlaThrSer 95
 Db 286 CGCAGATTTGCTCTCCACAGAGCCCGCTGACCTGGAACGCGATCAGAGATCCACCAAC 345
 Qy 96 TyrProAsnLeuCyLeuGlnAsnSerGluTTPLeuLeuAspGlnHisMet---Leu 114
 Db 346 TAACCACTTTCGTGCTTCAAGAAATCTGGGGAACAGATATATATCCATGATCACTTAAC 405
 Qy 115 LysValHisTyrProLysPheGlyValSerGluAspCyLeuTyrLeuAsnIleTyrAla 134
 Db 406 CGGAACACAGTAAATAGCTGATTTCAAGTAGAGATGCTTACTGAATGTATATGCC 465
 Qy 135 ProAlaHisAlaAspThrGlySerLysLeuProValLeuValTyrPheProGlyGlyAla 154
 Db 466 CCAAGTACTGCGCCCTGGGGCCCTCTGCTGCGACAGATGATGTTGGTTTCTGAGAGTGGC 525
 Qy 155 PheLysThrGlySerLysSerLysPheAspGlySerAlaLeuAlaIleTyrLysAspVal 174
 Db 526 TTCTCTGCGCGTTGCGCTTCAACCTACGAAAGCTCAGAGCTGAGGCGCGTGGGAAAGTG 585
 Qy 175 LeuValValValIleGlnTyrArgLeuGlyIlePheGlyPhePheThrTyrAspGln 194
 Db 586 GTGCTGCTGTTTCTGACAGTACAGCTGGGATCTCGGCTTTTCAAGCACCGGCAACAGC 645
 Qy 195 HisAlaProGlyAsnTTPAlaPheLysAspGlnValAlaIleLeuSerTyrValGlnLys 214
 Db 646 CATGCCCGCGGAAATGGGGGCTGCTGCGACAGATAGCTTTCGCGGTGGTGAAGAG 705
 Qy 215 AsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGluSerLysGly 234
 Db 706 AACATCGAAGCCTTTGGTGAAGACCGGACAGTGAACCTATTTGGTCAAGTGGCGGGA 765
 Qy 235 AlaIleSerValSerSerLeuLeuSerProMetAlaLysGlyLeuPheHisIleVala 254
 Db 766 GCGATGACGCTTCAAGACTAGTATATCAACCTTCAAGCCAGGCTGTGTTCAATCAAGCC 825
 Qy 255 IleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyrGluLysSer 274
 Db 826 ATCTCCAGAGTGTATGCGGGTCTGAAAGCCTTCACTACCTCCGATCCACTGACGTCA 885
 Qy 275 GluAspLeuGlnValAlaHisPheCyGlyAsnAsnAlaSerAspSerGluAlaLeu 294
 Db 886 -----GCCAAGACAGATCGCTACCTGCTGCTGACACAAATAGCAAGAGGTCAAG 939
 Qy 295 LeuArgCyLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGlnLysThrLys 314
 Db 940 GTGGAATGCTAGAGTCTCTTCAAGAGAGGTGAGCCTTGTTCACAGAGATGTCA 999
 Qy 315 SerPhe-----ThrArgVal 319
 Db 1000 TTCTTTGAGGCTACTCCATTAAGATCCAAAGAGATGTTGGTTCCTGAGCCCAAGTG 1059
 Qy 320 ValArgGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLys 339
 Db 1060 GTGATGTGTGTGTTCTCCAGAAATCTGTGGTCTTCAACCGGTGGGAGGTTAA 1119
 Qy 340 AlaIleProSerIleIleGlyValAsnAsnHisGlu-----CyGlyPheLeu 355
 Db 1120 CCTGTGCTTACTCTCTAGTGTTCACACGACAGATTCGATGGAATTAATCTTCTTC 1179
 Qy 356 LeuProMetLysGlyAlaProGluLeuSerGlySerAsnLysSerLeuAlaLeuHis 375
 Db 1180 ATGAAGATTACAGCTTAATCAGGTATGATGAACCAACCAATCATCAAAATATCTCTGG 1239

Qy 376 LeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisLeuValAlaAsnGluTyr 395
 Db 1240 AGTTTCACAGCTTCTGTGATATATCACTTAAGACAGAGTCCAGTGTGGCAGAGAGTAT 1299
 Qy 396 PheHisAsp-----LysHisSerLeuThrGluIleArgAspSerLeuLeuAspLeu 412
 Db 1300 TTGAATGACGCTATCATGATTAACCATGATGGAAGATGTTCCGAAATCGCTGTATGACCTTA 1359
 Qy 413 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAspAla 432
 Db 1360 ATCGAGAGATGCCACCTTTGTCTTACGCCACATCTCCAAAGCTGCGAGGTACCAAGATGCT 1419
 Qy 433 G1YAlaProValTyrPheTyrGluPheArgHisArgProGlnCyPheGluAspThrLys 452
 Db 1420 GGCTCCCGCTATCTGTATGACAGTTCAAGACACATGCTTCTTGGGACATCATCAATA 1479
 Qy 453 ProAlaPheValValAlaAspHisAlaAspGluValArgPheValPheGlyGlyAlaPhe 472
 Db 1480 CCT---TCAATGAGACAGACCAAGAGAGCTTTCCTATCTCTTGGCAGCCCTTTC 1536
 Qy 473 LeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluLysLeuLeuSerArg 492
 Db 1537 TCCAAAGCC-----TCATCTGACGTTGAGGAAGGAATTCAGCTTC 1578
 Qy 493 LysMetMetLysTyrTTPAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeu 512
 Db 1579 CAGATGATGAATATCTGGGCCCACTTGTCTCACTCGGAAACCCAAATGCGAAAGCTG 1638
 Qy 513 SerLeuTTPProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSer 532
 Db 1639 CCTACTGGCCACGCTTGGACAAAGATGAGAAATGACCTGCAAGTTGCACTTCGATCAAGA 1698
 Qy 533 LeuGlyGlnArgLeuLysGluProArgValAspPheTTPThrSerThrIleProLeuIle 552
 Db 1699 GTGGGTGTGAAGCTCAAGAGAAAGATGCTTTCGTGAG-TCAGCTTCGCAATC--- 1754
 Qy 553 LeuSerAlaSerAspMetLeuHisSerPro 562
 Db 1755 -----TCAGAAACCTTAGAGACACCA 1775
 RESULT 11
 AK078879 1944 bp mRNA linear HTC 03-APR-2004
 LOCUS AK078879
 DEFINITION Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030604P03 product:carboxylesterase 3, full insert sequence.
 ACCESSION AK078879
 VERSION AK078879.1 GI:26347580
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 Carrinci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL MEDLINE 99279253
 PUBMED 10349636
 2
 Carrinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtractions of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL MEDLINE 20499374
 PUBMED 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carrinci, P., Komoto, H., Akiyama, J., Nishi, K., Kitsuana, T., Teshiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,


```

Oy 269 HisaPRTYGLVysSergLusApleuGlnValAlaHisPheCySgLYAsnAanla 288
Db 885 ACAGATGTAAGCCCATTCCTGCTG-----GTTGCTACTTTCTTGCGGTAAACT 888
Oy 289 SerAapSergLusAlaLeuLeuAryCySleuAgtThrLyPProSerLySgLYleuLeuThr 308
Db 889 ACCAATCAGCTGTATATGTTATGCTCCGCCAGAAAGACAGAGATGMACTACTGAG 948
Oy 309 LeuSergLusLyThrLySerPhe----- 316
Db 949 ACCTCACTAAATTTGAATCTTTTAACTGCACTTACTTGAATCCAAAGAGACTAT 1008
Oy 317 -----ThrAValValAspGlyAlaPhePheProAsnGluProleuApleuLeu 333
Db 1009 CCCTTCCTCCCTACGCTATGATGAGTTCGCGCAAGGACACAGAAAGATCTCG 1068
Oy 334 SerGlnLyAlaPheLyAlaLeuProSerLlellegLYAlaAsnAanleSgLYsGly 353
Db 1069 GCTGAGAAAGATTTCAGCAGCTGCTCCCTACATAGTGGCATCAACAGCAGAGATTGGC 1128
Oy 354 PheLeuLeuPrometLySgLYAlaProGluLeuSergLYSer---AsnLySerLeu 372
Db 1129 TGGATCATTCACAACTTATGCGCTATCCACTGCTGAAGCAAACTGAGCAGAAACA 1188
Oy 373 AlaLeuHisLeuLleGlnAsn-----LleuHisSleProProGlnLyThrHis 389
Db 1189 GCCATTCCTCTCTTGTGAAGTCCACCAACACTTAATCTCGAAGATGATGATCGA 1248
Oy 390 LeuValAlaAsnGlnLyThrPheHisAplyHisSergLYLeuGlnLyArgAapSerLeu 409
Db 1249 GTGGTGGCTGAGAACTATTAGAGAGACAGATGACCTCACAAAGAAAGACCTGTTTC 1308
Oy 410 LeuAapLeuLeuGlyAspValAlaPhePheValValProAlaLeuLleThrAlaArgThrs 429
Db 1309 CAGCATGTAGTGGTGAATGTTGTTGTTCCATCCATGATTTGTCTCGAAGTCC 1368
Oy 430 ArgAapAlaGlyAlaProValLyThrPheLyGlnPheAsnGlyArgProGlnCySapPheGlu 449
Db 1369 AGAGATGCTGAGAGCTCCACCTATATGATGATTGATGATCGCCCAAGCTTGTATCG 1428
Oy 450 AspThrLySProAlaPheValLyAlaAsnHisAapSgLYValAlaArgPheValPheGly 469
Db 1429 GCCATGAGACCCAGCAGATATAGAGACCATGTGATGATGATCTTCACATATTGGA 1488
Oy 470 GlyAlaPheLeuLeuGlyAspLleValMetPheGlnGlyAlaThGlnGlnGlnLySleu 489
Db 1489 TCTCCATTTTAAA-----GATGGTGGCTCAGAAAGAGAGACCAAC 1530
Oy 490 LeuSergLySLeuMetLySerLyThrAlaThrPheAlaArgThrsGlyAsnProAsnGly 509
Db 1531 CTCAGCAAGATGTATGATAATTTCTGGGCCAATTTGCTCGGAATGGGAACCCCAATGCT 1590
Oy 510 AsnAapLeuSergLyThrProAlaLyAsnLeuThrGlnGlnLySleuApleu 529
Db 1591 GGAGGGCTGGCCCACTGGCCAGAAATATGACAGAAAGAGGATATCTGAAGATGGTGGC 1650
Oy 530 AsnMetSerLeuGlnLyAlaArgLeuLySgLYProAlaGlyAlaPhePheThrTherThrle 549
Db 1651 TCAATTCAGGCAAGCCAGAGGCTGAAAGGACAAAGAAAGAGATTGTTGGCTGAG----- 1704
Oy 550 ProLeuLleLeuSergLySLeuAapMetLeuHisSerPro 562
Db 1705 -----CTCAGGGGCCAAGAGATCAGGCCCAGAGGCCA 1734

```

```

RESULT 12
CF111083 1918 bp mRNA linear EST 23-JUL-2003
LOCUS Shultzomiza0434 Rat lung airway and parenchyma cDNA libraries
DEFINITION Rattus norvegicus cDNA clone Contig345 5', mRNA sequence.
ACCESSION CF111083
VERSION CF111083.1 GI:33167595
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)

```

```

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1918)
AUTHORS Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
Padua,A.M., Gureke,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
Plopper,C.G. and Buckpitt,A.R.
TITLE Gene expression analysis in response to lung toxicants: I.
Sequencing and microarray development
JOURNAL Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
COMMENT Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mshultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 1918.
FEATURES
source location/Qualifiers
1..1918
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="Contig345"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/note="Organ: lung; Vector: pGSM-112f(-); Site 1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."
ORIGIN
Alignment Scores:
Pred. No.: 1,286-106 Length: 1918
Score: 1097.50 Matches: 239
Percent Similarity: 55.82% Conservative: 87
Best Local Similarity: 40.92% Mismatches: 207
Query Match: 35.64% Indels: 52
DB: 7 Gaps: 9
US-10-023-515-2 (1-581) x CF111083 (1-1918)
Oy 5 LeuThrsSerSerAlaSergLntPrCySapPheLeuLleuGlnProleuGlnLyHis 24
Db 27 CTACAAATGGCTCCTACCTCGCTGCTGCTTTCTTCTGCGCGACAGCTTGG--- 83
Oy 25 ArgGlnTrpGlyLyThrGlyProSeraLagGlyProGlnAAsnThrArgLeuGly 44
Db 84 -----GGTAT--CCATCCTCACCACTGTGGTGACACTGTTAAAGCC 124
Oy 45 TrpLleGlnLySgLYAlaThrValLeuGlySerProValAsnValPheLeu 64
Db 125 AAAGTCTGGGAGATATGATCAATTGGAAGATTGGACAGCCTGTGGCTTTCTCG 184
Oy 65 GlyValProPheAlaAlaProProleuGlySergLeuAapPheThrAsnProGlnProAla 84
Db 185 GGAATCCCTTCGCCAATGCCCTTGTGGCTCTTGAGAGTTTGCTCCACACAGCTGCA 244
Oy 85 SerProLTPAapAenLeuAryGlnAlaThrseryTrProAenLeuCyLeuGlnAsn--- 103
Db 245 GAGTCTTGAACCTTGTGAAGAAATACTACTCTTACCCACCTATGCTCTCAAGATCT 304
Oy 104 -----SergLntPrLeuLeuAapGlnHisMetLeuLyValHis 117
Db 305 GTTGAGGGCAGGTTCTTCAGAGCTTTTTCACCAACAGAAAGAAACATTCCTTTACG 364
Oy 118 TyrProLySphGlyValSergLusApleuLySleuAenLleLyAlaProAlaHis 137

```

Db 365 TTT-----TTGAAGATGCGCTTAACCTGAATGTTTAATCTCCCGCTGAC 409
 Qy 138 AlaSerThrGlySerIleuProValIleuValIleuProGlyGlyAlaPheIleThr 157
 Db 410 TTGCAAAAGAACACCGCGTTTACAGTAGTGGTGTGATTCATGAGAGGTGACGTGTAGT 469
 Qy 158 GlySerAlaSerIlePheAspGlySerAlaIleuAlaIleuGlyGlyAspValIleuVal 177
 Db 470 GGTGGATCATCTCCACTTAATGATGACAGCGCTCTCTCCCATGAATAATGGTGGGTG 529
 Qy 178 ValValGlnTyrGlnIleuGlyIlePheGlyPhePheThrIleuArgGlnIleAlaPro 197
 Db 530 ACCATTGATGATGCGCTTGGCATCTGGGAGATTCTTGACGACAGGGATGACACACCG 589
 Qy 198 GlyAsnThrIlePheIleuAspGlnValAlaIleuSerTyrValGlnIleuValIleuGly 217
 Db 590 GGGACATCGGGGTCAATTGGACACAGCTGGTGCACCTGCACAGGGGTCCAGACAAATATGG 649
 Qy 218 PhePheGlyGlyAspProSerSerValIlePheGlyGlySerAlaGlyAlaIleSer 237
 Db 650 AACTTTGGGGGGACACCGAGCTGTGTGACCATCTTTGGAGAACTGACAGAGTTCACT 709
 Qy 238 ValSerSerIleuIleuSerProMetAlaIleGlyIleuPheIleuValIleuGly 257
 Db 710 GTCTCTGTTCTGTGGTATCTCTCTGGCCAGAAACCTGTCACAGGCGCATTTGTGAG 769
 Qy 258 SerGlyValAlaIleIleProGlyIleuGlnAlaIleAspTyrGlnIleuSerGlyIleu 277
 Db 770 AGTGGGTGTCCTCAATCTGTCTGTATTCACACAGATAGCAGCCCATCTGAATCTG 829
 Qy 278 GlnValValAlaIlePheCysGlyIleuAsnAlaSerAspSerGlyIleuIleuArgCys 297
 Db 830 -----ATTGCTACTCTTCTGGGTGTAAACACACACATCAGCTGTATGTCATTCG 883
 Qy 298 LeuArgThrIleProSerIleGlyIleuIleuIleuSerGlnIleuThrIleuSerPhe 316
 Db 884 CTGGCCGACAGACAGAGATGATCTCTGACACTCTTAATAATGAACTTTTAA 943
 Qy 317 -----ThiArgValAlaArgGly 322
 Db 944 CTGACCTATTGGAAACCAAAAGAGAGTATCCCTCTCACTGATGTATGTATGTA 1009
 Qy 323 AlaPheProAsnGlnProIleuPheIleuSerGlnIleuValPheIleuAlaIlePro 342
 Db 1004 GTGGTGTCTCAAAAGACACAAAGAGATCTGGCTGTAGAAAGTTTCAACGCGTCC 1063
 Qy 343 SerIleIleGlyValAsnAsnIleGlyCysGlyPheIleuProMetIleGlyIleuAlaPro 362
 Db 1064 TACATAGTGGGCATCAACAGACAGATTTGGCTGATCTTCAATCTTATGGCGTAT 1123
 Qy 363 GlnIleuSerGlySer---AsnIleSerIleuAlaIleuIleuIleGlnAsn----- 379
 Db 1124 CCACTCTCCGAAGGAAACCTGACCCAGAAACAGCAATCTCCCTGTGGAAAGTCTCA 1183
 Qy 380 ---IleuIleIleProGlnIleuIleuIleuValAlaAsnGlnTyrPheIleAsp 398
 Db 1184 CCAACACTGAATAATCTGAGAAATATGTTCCAAAGCGCTTGTGAAAGATTTTGAAGCG 1243
 Qy 399 LysIleAspIleuThrGlnIleuArgAspSerIleuAspIleuIleuGlyAspValPhePhe 418
 Db 1244 ACAGATGACCTCGCGAAAGAAAGACCTGTTCACAGGCTTGGTTCAGATGTATATT 1303
 Qy 419 ValValProAlaIleuIleuThrAlaArgTyrIleuArgAspAlaGlyAlaProValTyrPhe 438
 Db 1304 GGTCTCCCATAGTGAATGTGTCTGAAAGTACAGAGATGTGGAGCCCCCACTTATG 1363
 Qy 439 TyrGlnPheArgIleArgProGlnCysPheGlyAspThrIleuProAlaPheValIleuAla 458
 Db 1364 TATGAATTTGAGTATCGCCCAAGCTTTGTATCAGACCATGAGGCCCAAGACAGTATGGA 1423
 Qy 459 AspIleIleAspGlyIleuArgPheValPheGlyGlyAlaPheIleuIleuValIleuAla 478
 Db 1424 GACCATGTATAGAACTTTCTCAATTTTGGATCTCCATTTTAAAA----- 1471

[illegible]

30 ThrGlyProserAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPheGlnGlyLys 49

```

Db      84  ACTGGGCCCCGAAGTTCTCAGCCTGAGTACAGACACCCCTGGGTCTGTGGAGGCCG 143
Qy      50  GlnValThrValLeuGlySerProValProValaenValPheLeuGlyValaProPheAla 69
Db      144  CAGGTGGGGGTAAAGGACAGACCGCCTTGTAAGTCTTCTTGCGCATTCATTCATTTGCC 203
Qy      70  AlaProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrrAspAsn 89
Db      204  CAGCGGCGCATCGGGCCTGACCGGTTCTGACGCCACACCCAGCAGACGCTGGAGGGGT 263
Qy      90  LeuArgGluAlaThrSerTyrProaenLeuGlyLeuGlnAsnSerGluTrpLeuLeuLeu 109
Db      264  GTGGCGGAGTCCAGACCTGCGGCCCAATGTCTTACAGACGATGAGACATGACACAGC 323
Qy      110  AspGlnHisMetLeuLysValHisTyrProLysPheGlyValSerGluAspGlyLeuTyr 129
Db      334  AGCAAGATTTGCTCTCAACGGAAACAGACAGATCTTCTCCGTTTCAGAGAGATCGCTGCTC 383
Qy      130  LeuAsnIleTyrAlaProAlaHisAlaAspThrGlySerTyrLeuProValIleuValTrr 149
Db      384  CTCMACTGTATAGCCCAAGCTAGAGTCCCGCAGGGTCCGTAAGCCGCTCATGGTATGG 443
Qy      150  PheProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 169
Db      444  GTCCATGAGAGCGCTCTGATACCTGGCGCTGCACCTCTTACATGATCAGATCAGCTGGCT 503
Qy      170  AlaTyrGluAspValLeuValValValGlnTyrArgLeuGlyTyrLeuPheGlyPhePhe 189
Db      504  GCGTATGGGAGATGTGCTGTGTACAGTACAGTACCGGCTTGGGCTCTTGGCTTCTTTC 563
Qy      190  ThrThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeu 209
Db      564  AGCATCTGAGATGAGATCATCTGACACCGAGGCTTCTTGAATGTGTAGTCTGTTTG 623
Qy      210  SerTrrValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePhe 229
Db      624  CGCTGGGTGCAAGAAACATCGCCCTTCGGGGGTGACCTCACTGATGTCATCTTCTT 683
Qy      230  GlyGluSerAlaGlyValaIleSerValSerSerLeuIleLeuSerProMetAlaLysGly 249
Db      684  GGTGATCTGCGCGTGGAGCATCTCTGCGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTG 743
Qy      250  LeuPheHisLysValIleMetGlnSerGlyValAlaIleIlePro---TyrLeuGluAla 268
Db      744  CTGTTTCCACAGGCCATCCACAGAGTGGGTGATCATCACCCAGGATCATTCAGACTTT 803
Qy      269  HisAspTyrGluLysSerGluAspLeuGlnValAlaHisPheCysGlyLysAsnAla 288
Db      804  CACCTTGGCCCCCTAGCT-----CAGAAATCGGAAACACCTTGGCTGACGCTCC 854
Qy      289  SerAspSerGluAlaLeuLeuArgCysLeuArgThrLysProSerTyrGlyLeuLeuThr 308
Db      855  AGCTCCCCCGCTGAGATGTGTGACGTGCTTTCAGAGAAAGAGAGAGAGAGAGAGCTGTGC 911
Qy      309  LeuSerGlnLysThrLysSer-----PheThrArgValValAlaAspGlyAlaLaphePro 326
Db      912  CTTACCAAGAAAGCTGAAATATCTATCTCTTCCACCGTGAATGACCTGTCTTCCCC 971
Qy      337  AsnGluProLeuAspLeuLeuSerGlnLysValaPheLysAlaIleProSerIleIleGly 346
Db      972  AAAAGCCCAAGAGAACTCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1031
Qy      347  ValaAsnAsnHisGlyCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSer 366
Db      1032  GTCAACAAACCAATGAGTTCAGCTGCTCATCCAG----- 1066
Qy      367  GlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGln 386
Db      1066 ----- 1066
Qy      387  TyrLeuHisLysLeuValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArg 406
Db      1066 ----- 1066

```

```

Qy      407  AspSerLeuLeuAspLeuLeuGlyAspValPhePheValIleProAlaLeuIleThrAla 426
Db      1066 ----- 1066
Qy      427  ArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGln 446
Db      1067 -----GATTCGTGAAGCCCTGTCTTTCTATGATTCAGATCAGACCAAGT 1114
Qy      447  CysPheGluAspThrLysPheProAlaPheValIleValaAsnHisAlaAspGluValaArgPhe 466
Db      1115  TCTTTTGGAGATCAAACTCTGCTGTGGTGAAGGTGTGATCAAGGGGCCAGGGTCTTTT 1174
Qy      467  ValPheGlyGlyAlaPheLeu-----LysGlyAspIleValMetPheGluGlyAlaThr 484
Db      1175  GTTTTCGAGAGTCCCTTCTCTTACAGACAGAGCTCCCGCTGCTTTCAGAGGCCACA 1234
Qy      485  GluGluGluLysLeuLeuSerArgLysMetLysTyrTrrPalAlaThrPheAlaArgThr 504
Db      1235  GAGAGAGAGAGACACTTAAGCTTCAACATGATGGCCAGTGGACCACTTCCCGGACA 1294
Qy      505  GlyAsnProAsnGlyLysAsnAspLeuSerLeuTrrProAlaTyrAsnLeuThrGluGlnTyr 524
Db      1295  GGGAGACCCCAATAGCAAGGCTCTGCTCTTGGCCCAATTCAACAGCGGAGCAAAATAT 1354
Qy      525  LeuGlnLeuAspLeuAsnMetSerLeuGlyHisArgLeuGluProArgValaAspPhe 544
Db      1355  CTGGAGATCAACCCAGCTGCGCACGGGCCGAGACAGAAATTCAGAGAGGCTTGATGCATTTC 1414
Qy      545  TrrThrSerThrIlePro 550
Db      1415  TGGTCAGAGAGCGCTCCCC 1432

```

RESULT 14

```

CRE61368      1859 bp      mRNA      linear      HTC 18-AUG-2004
LOCUS      Tetradon nigroviridis full-length cDNA.
DEFINITION      CRE61368
ACCESSION      CRE61368
VERSION      CR61368.1 GI:51137813
KEYWORDS      HTC; cDNA; full-length; Tetradon nigroviridis.
SOURCE      Tetradon nigroviridis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoleleostei;
      Tetraodontidae; Tetraodontidae; Tetradon.
REFERENCE      1 (bases 1 to 1859)
AUTHORS      Genoscope.
TITLES      Direct Submision
JOURNAL      Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
      : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
      (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      The sequences are based on single pass reads.
      More information available at
      http://www.genoscope.cns.fr/tetradon.

```

FEATURES

```

source      1..1859
      /organism="Tetradon nigroviridis"
      /mol_type="mRNA"
      /db_xref="taxon:99883"
      /cissue_type="Liver"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      4.67e-102      Length:      1859
Score:      1055.00      Matches:      228
Percent Similarity:      59.70%      Conservative:      89
Best Local Similarity:      42.94%      Mismatches:      184
Query Match:      34.26%      Indels:      30
DB:      3      Gaps:      10

```

US-10-023-515-2 (1-581) x CR61368 (1-1859)

```

Qy 37 ProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySer 56
Db 113 CTTAAATCCATACCAAACTGCGACGCTGAGAGTGATATCAAGACGGGAAGGAAA 172
Qy 57 ProValProValAlaPheLeuGlyValProPheAlaAlaProPheLeuGly---Ser 75
Db 173 GACCGTGGAGTCCATGATACCTGGGTGTGCTCTTCTCTAAGCACTCGGGTCCGCT 232
Qy 76 LeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArgGlyAlaThrSer 95
Db 233 CTGAGCTGGCTCCACCCGACCTGTAGAGGCTTGGAGAGAGAGATGCTACCAAG 292
Qy 96 TyrProAsnLeuGlyLeuGlnAsnSerGlu-----TrpLeuLeuAspGlnHisMet 113
Db 293 CAACCGGTATGTGTGTTCAGGATTTAGAGTTTGACAAACAATATATTAAGAACTTAAT 352
Qy 114 LeuLeuValHisTyrProLysPheGlyValSerGluAspGlyLeuTyrLeuAsnIleTyr 133
Db 353 TTAGAAATCAACCTTCA-----GACATTTCAAGAAAGCTCTTACCTCAACATTTAC 406
Qy 134 AlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValThrPheProGlyGly 153
Db 407 ACTCTCGCCACACAGACCTGAGAAATACCAAGCTCCAGTTATGTGTCATCCACGGTGA 466
Qy 154 AlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleTyrGluAsp 173
Db 467 GGGTGTGTGTAGGTCGGCTTCAATGATGATAGTGCTGCCCTGGCTGCTTACCAAGT 526
Qy 174 ValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThrTrpAsp 193
Db 527 GTTGTGTGTGTGTATCCAGTACGATCGAGAAATCTGGGCTTCTGAGCACTGTGTAT 586
Qy 194 GlnHisAlaProGlyValAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrpValGln 213
Db 587 AAGCAATCCAGCAAAACATGTGCTTTCGATCAAGCTCAAGCTCTGAAAGGGTCCAG 646
Qy 214 LysAsnIleGlnPhePheGlyLysAspProSerSerValThrIlePheGlyLysSerAla 233
Db 647 GAGCAATTCACAACTTGTGGAGAGACCCAGATTTAGTACCAATTTGGCAGAGCTGCT 706
Qy 234 GlyAlaIleSerValSerSerLeuLeuSerProMetAlaLysGlyLeuPheHisLys 253
Db 707 GGTGAGTGAAGTATCTCTCTGCTGCTCACTGTCTGAGGGGCTTTTCCGCCAC 766
Qy 254 AlaIleMetGlySerGlyValAlaIleLeuProTyrLeuGlyAlaHisAspTyrGluLys 273
Db 767 GCAATTTGCTGAAGAGTGCACCTGTCGAATG---CATATATTTGCTCAAGAT---GACCC 820
Qy 274 SerGluAspLeuGlnValAlaHisPheGlyAspAsnAlaSerAspSerGluAla 293
Db 821 CAGCCAGTGTACAACTTTGTGCAAAATGATCTGGTTCACATGAAAGCAGAAAG 880
Qy 294 LeuLeuArgGlyLeuArgThrLysProSerLysGlyLeuLeuThrLeuSer---GlnLys 312
Db 881 ACTGCTGATTCATCAAGACCTGAGACGTGAAGAACATTTTACCTTTTACCTAAGAAA 940
Qy 313 ThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAspLeu 332
Db 941 GAGTTAAGGATCTTAAATACACGATGATATTTCTCAAAAATGACAGAGAAATGTG 1000
Qy 333 LeuSerGlnLysAlaPheLysAlaIleProSerIleLeuValAlaAsnHisGlnLys 352
Db 1001 TTCTTGAAGCAATGAACTTTCACCTGCTTTCATGACTGAGTCAATTAATGAAGGT 1060
Qy 353 GlyPheLeuLeuProMetLysGlyAlaPro-----Gln 363
Db 1061 GGTGTGTACTTCTCGGTACTTGTCTCTCAAAATGACAGAGGGGATGACAGAA 1120
Qy 364 IleLeuSerGlySerAlaLysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIle 383
Db 1121 CAGCTCTCTGCT-----TTCTGTCTGTGTCTATCTCT 1153
Qy 384 ProPro-----GlnTyrLeuHisLeuValAlaAsnGlnTyrPheHisAspLysHis 400

```

```

Db 1154 GATCCCAAGATGACAGCTTAGAAAACGGATTTCTTAATGATATATACAGAACTGGGAG 1213
Qy 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
Db 1214 GATCTGTGAATAACAGACGGGTACACTGAGGTGCTTGGAGACATGATATTTGTCAATT 1273
Qy 421 ProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGlu 440
Db 1274 CTGCCATTAAAGGCTGCTATGCTCAGAGATGAGGTCTCTCTGTACTCTGTATGAG 1333
Qy 441 PheArgHisArgProGlnCyPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
Db 1334 TTTCAGCATTCCTCCCAAGTTTCTTAAAGATTAAGATCAAGCTTGTATGAGAGTACCAT 1393
Qy 461 AlaAspGluValArgPheValPheGlyGlyLysAlaPheLeuLysGlyAspIleValMetPhe 480
Db 1394 GGAATGAAATATTTTGTGTTTGGGTTTTCACAGTTTCTCACATCAATCTAAGC 1453
Qy 481 GluGlyAlaThrGluGluGluLysLeuSerArgLysMetLysTyrTyrAlaThr 500
Db 1454 AGTGGTGCACCTGAAGAAAGAGACAGATTGACCAATCATGATGAGCTACTGGGCAAC 1513
Qy 501 PheAlaArgThrGlyAsnProAlaGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeu 520
Db 1514 TTTGCTGCACAGGAGTCTCTTATGACAGGGTCTTGTCCACTGGCCAAAGTATGAGAA 1573
Qy 521 ThrGluGlnTyrLeuGlnLeuAspLeu---AsnMetSerLeuGlyGlnArgLeuGlu 539
Db 1574 ACGAGAGAGTACTCGGGAATGATGCAAAAGAGCAAGATTTTATCAAGGTTTGAAAAA 1633
Qy 540 ProArgValAspPheTrpThrSerThrIlePro 550
Db 1634 GACAAATATCTCTCTTACCACCAAGCCCCCA 1666

RESULT 15
CR634885 1817 bp mRNA linear HTC 18-AUG-2004
LOCUS
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR634885
VERSION CR634885.1 GI:51131330
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorphia; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1817)
REFERENCE
Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
2 rue Gaston Crémieux, CP 5706 - 91057 Evry cedex - FRANCE
COMMENT (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source
1..1817
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Alignment Scores:
Pred. No.: 2,566-101 Length: 1817
Score: 1048.00 Matches: 227
Percent Similarity: 59.51% Conservative: 89
Best Local Similarity: 42.75% Mismatch: 185
Query Match: 34.04% Indels: 30
DB: 3 Gaps: 10

```

US-10-023-515-2 (1-581) x CR634885 (1-1817)

```

Qy 37 ProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValGlnValThrValLeuGlySer 56
Db 93 CCGAAATTCATCAACACTGGCAGCTGAGAGGATATCATGAGCTGAAAGGGGAAA 152
Qy 57 ProValProValAsnValPheLeuGlyValProPheAlaAlaProProLeuGlySer 75
Db 153 GACGCTGAGATCATGACATACCTGGGTGGTGGCTTTGCTAACACCTCTGGTCTGCT 212
Qy 76 LeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArgGlyAlaThrSer 95
Db 213 CTGAGGCTGGCTCCACCCAGCTGTAGAGCTTGGAAAGAGTGAAGATGCTACCAAG 272
Qy 96 TyrProAsnLeuCysLeuGlnAspSerGlu-----TrpLeuLeuLeuAspGlnHisMet 113
Db 273 CAACCGCTGATGTGTTCAGAGATTAGAGTTGCAACAACAATTAATTAACAACCTTAAT 332
Qy 114 LeuValValHisThrProLysPheGlyValSerGluAspCysLeuTyrLeuAsnIleTyr 133
Db 333 TTAGAAGTCACACTTCA-----GACATTTCAAGAGAGCTGCTTACTCAACATTTAC 386
Qy 134 AlaProAlaHisAlaAspThrGlySerTyrLeuProValLeuValTrpPheProGlyGly 153
Db 387 ACTCTGCAACAGACAGCTGAGAAATACCAAGCTCCAGATTATGCTGATCCAGGTGGA 446
Qy 154 AlaPheLysThrGlySerIleAspIlePheAspGlySerAlaLeuAlaAlaTyrGluAsp 173
Db 447 GGGTTTGTATTAGTTGGCTTCATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506
Qy 174 ValLeuValValValValIleGlnTyrArgLeuGlyIlePheGlyPhePheThrThrTrpAsp 193
Db 507 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 566
Qy 194 GlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrpValGln 213
Db 567 AAGCATATGTCAGAGAAACATTGCTCTGATCAGTCCAGGCTGTAAGAGGCTCCAG 626
Qy 214 LysAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGlySerAla 233
Db 627 GAGCATATTCACAACTTGGAGAGAGACCAATTAGTACCAATTTGGCAAGCTGCT 686
Qy 234 GlyAlaIleSerValSerSerLeuLeuLeuSerProMetAlaLysGlyLeuPheHisLys 253
Db 687 GGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 746
Qy 254 AlaIleMetGlySerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyrGlyLys 273
Db 747 GCAATGCTGAAAGTGGCACTGCTCAATG---CATATATTGCTCAAGAT---GACCCC 800
Qy 274 SerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSerGluAla 293
Db 801 CAGCCAGTGTAAACATTTGTGCGCAATGATCTGTTGAGCATGAAACACAGAAAAG 860
Qy 294 LeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer---GlnLys 312
Db 861 ACTGCTGATTCATAGAACTGGAAGCTGGAAACCATTTTACCTTTAAGGAAAGAAA 920
Qy 313 ThrLysSerPheThrArgValValAlaAspGlyAlaPhePheProAsnGluProLeuAspLeu 332
Db 921 GAGTTAAGTATCTTATATACGATGATATTTCTCAAAAATTCAGACAGAAAGTTG 980
Qy 333 LeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAlaAsnHisGluCys 352
Db 981 TTCCTTGAGCATGAACCTTCTCACTGTGCTTTCATGACTGAGTCAATATATATGAGGT 1040
Qy 353 GlyPheLeuLeuProMetLysGluAlaPro-----Gln 363
Db 1041 GGTGTGTTACTTGTCTGCGTACTTGTCTCTCCAACTGACAGAGGAGTGAACAGAA 1100
Qy 364 IleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIle 383
Db 1101 CACGTCCTGGT-----TTCCTGTCTGTGTCTATCT 1133

```

```

Qy 384 ProPro-----GlnTyrLeuHisLeuValAlaAsnGluTyrPheHisAspLysHis 400
Db 1134 GATCCCAAGATGACAGCTTATGAAAACGATTTCTTAATGATATATACAGAACTGGGAG 1193
Qy 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
Db 1194 GATCTGTGAAAACAGAGAGCGGTACCTGAGGTGCTTGGAGACATGATATTTGTCAAT 1253
Qy 421 ProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGlu 440
Db 1254 CTGCAATTAAGCTGCTATTTGCTCAAGAGATGACAGGTGCTCTGTGTACCTGTATGAG 1313
Qy 441 PheArgHisArgProGlnCysPheGlyLysPheThrLysProAlaPheValLysAlaAspHis 460
Db 1314 TTGCAGCATTTCCCAAGTTTCTTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1373
Qy 461 AlaAspGluValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
Db 1374 GGAATGAAATATTTTGTATTTGGCTTTGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTT 1433
Qy 481 GluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyrTrpAlaThr 500
Db 1434 AGTGGTGCACCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1493
Qy 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeu 520
Db 1494 TTGCTCCACAGGGGTCTCTTATGACAGGGCTTTGTCACAGTGGCCAAAGTATGAGAA 1553
Qy 521 ThrGluGlnTyrLeuGlnLeuAspLeu---AsnMetSerLeuGlyGlnArgLeuLysGlu 539
Db 1554 ACGAGAGAGTACCTGGAATTTGATGCAAAAGAGCAAGATGTTTATCAGAGGTTTGAAGAA 1613
Qy 540 ProArgValAspPheTrpThrSerThrIlePro 550
Db 1614 GACAAATATCTCTCTTCAACCCAGACCCCCCA 1646

```

Search completed: June 15, 2005, 13:29:40
 Udb time : 5297 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 15, 2005, 09:44:41 ; Search time 283 Seconds
(without alignments)
3359.282 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079
Sequence: 1 MPQGLTSSASQWCFLLIQP.....PLSLTFLSLIQPFPPFCAP 581

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delgap 6.0 , Fgapext 7.0
Delgap 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/6664091/runat_14062005_133641_10487/app_query.fasta.1.775
-DB=Issued_Patents_NA -QFMT=fastcap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=6664091_0CGN_1_1_69_0runat_14062005_133641_10487 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3079	100.0	1746	4	US-10-023-515-3	Sequence 3, Appl1
2	3079	100.0	2158	4	US-10-023-515-1	Sequence 1, Appl1
3	1230	39.9	2117	4	US-09-949-016-3799	Sequence 3799, Ap
4	1230	39.9	2169	4	US-09-949-016-555	Sequence 555, App
5	1230	39.9	2191	4	US-09-595-682B-27	Sequence 27, Appl
6	1140.5	37.0	1717	4	US-09-595-682B-20	Sequence 20, Appl
7	1138.5	37.0	1701	3	US-09-264-737-3	Sequence 3, Appl1
8	1129.5	35.7	965	4	US-09-799-451-155	Sequence 155, App
9	793	25.6	2802	4	US-09-949-016-2515	Sequence 2515, Ap
10	789.5	25.6	4960	4	US-09-949-016-1017	Sequence 1017, Ap
11	782	25.4	2862	4	US-09-949-016-2516	Sequence 2516, Ap
12	775	25.2	4233	4	US-09-491-356C-5	Sequence 5, Appl1

13	764.5	24.8	1845	1	US-07-732-962A-1	Sequence 1, Appl1
14	764.5	24.8	1845	5	PCT-US92-06106-1	Sequence 1, Appl1
15	764.5	24.8	2158	4	US-09-949-016-1192	Sequence 1192, Ap
16	764.5	24.8	2158	4	US-09-949-016-1193	Sequence 1193, Ap
17	764.5	24.8	2256	2	US-08-318-826A-5	Sequence 5, Appl1
18	764.5	24.8	2256	2	US-08-370-156-1	Sequence 1, Appl1
19	764.5	24.8	2256	3	US-08-814-095-1	Sequence 1, Appl1
20	764.5	24.7	3096	2	US-08-318-826A-6	Sequence 6, Appl1
21	759.5	24.7	3096	2	US-08-370-156-3	Sequence 3, Appl1
22	759.5	24.7	3096	2	US-08-814-095-3	Sequence 3, Appl1
23	759.5	24.7	5767	4	US-09-810-861B-3	Sequence 3, Appl1
24	759.5	24.7	14446	4	US-09-810-861B-4	Sequence 4, Appl1
25	755.5	24.5	1725	4	US-09-810-861B-5	Sequence 5, Appl1
26	755.5	24.5	3016	2	US-08-318-826A-7	Sequence 7, Appl1
27	755.5	24.5	3016	2	US-08-370-156-5	Sequence 5, Appl1
28	755.5	24.5	3016	3	US-08-814-095-5	Sequence 5, Appl1
29	755	24.5	2400	6	5215909-13	Patent No. 5215909
30	755	24.5	2400	6	5215909-13	Patent No. 5215909
31	752	24.4	2381	2	US-08-318-826A-9	Sequence 9, Appl1
32	752	24.4	2416	2	US-08-318-826A-8	Sequence 8, Appl1
33	752	24.4	2416	3	US-09-334-489-1	Sequence 1, Appl1
34	752	24.4	2416	3	US-09-334-489-2	Sequence 2, Appl1
35	752	24.4	2444	4	US-09-949-016-5275	Sequence 5275, Ap
36	677	22.0	68667	4	US-09-949-016-17017	Sequence 17017, A
37	673.5	21.9	1764	4	US-09-491-356C-3	Sequence 3, Appl1
38	664.5	21.6	4436	4	US-09-491-356C-6	Sequence 6, Appl1
39	660	21.4	2184	1	US-08-445-050-8	Sequence 8, Appl1
40	660	21.4	2184	1	US-08-204-691-8	Sequence 8, Appl1
41	660	21.4	2344	3	US-09-347-878-31	Sequence 31, Appl1
42	660	21.4	2375	4	US-09-949-016-3976	Sequence 3976, Ap
43	660	21.4	2428	1	US-08-445-050-1	Sequence 1, Appl1
44	660	21.4	2428	1	US-08-204-691-1	Sequence 1, Appl1
45	660	21.4	2428	4	US-09-355-295B-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-10-023-515-3
Sequence 3, Application US/10023515
Patent No. 6664091
GENERAL INFORMATION:
APPLICANT: Curtis, Rory A. J.
APPLICANT: Silos-Santiago, Immaculada
TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
FILE REFERENCE: 10448-122001
CURRENT APPLICATION NUMBER: US/10/023,515
CURRENT FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,369
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/279,508
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1746
TYPE: DNA
ORGANISM: Homo sapiens
US-10-023-515-3
Alignment Scores:
Pred. No.: 0
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0
US-10-023-515-2 (1-581) x US-10-023-515-3 (1-1746)
1 MetProGInGlyLeuThrSerSerzLaserGInTPCySphePheLeu1leuGInPro 20
|||||

Db 1 ATGCCAGAGGACTTACTTACTGCTTCACATGATGCTTTTCTGATTCACAGCC 60
Qy 21 LeuLeuGlyHisArgGlnTTPGlyLyserThrGlyProSerAlaGluGlyProGlnArgAsn 40
Db 61 CTGTGGAGACACACAGTGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGAAC 120
Qy 41 ThrArgLeuGlyTPRIleGlnGlyLeuGlnValThrValLeuGlySerProValProVal 60
Db 121 ACCAGGCTGGGATGATTCAGGGGACAGCAAGCTCATGTGCTGGGAGCCCTGTGGCTGTG 180
Qy 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
Db 181 AAGGTGTCTCCGAGAGTCCCTTTGCTGCTCCCGCTGGAGTCCCTGGATTTAGAAC 240
Qy 81 ProGlnProAlaSerProTPAAsnLeuArgGluAlaThrSerTyProAsnLeuCyS 100
Db 241 CCGGAGCTGCATGCGCTGGAGTAATCTTGAGAAAGCCACCTTACCTTAATTTGTGTC 300
Qy 101 LeuGlnAsnSerGlyTPLeuLeuLeuAspGlnHisMetLeuValHisTyProLys 120
Db 301 CTCACAGAACTCAGATGGCTGCTTAAAGTCAACATGCTCAAGGTGCAATTCAGGAAA 360
Qy 121 PheGlyValSerGlyLeuArgLysLeuValLeuAsnIleTyValProAlaHisAlaPheThr 140
Db 361 TTGGAGGTGCAGAAAGATGCTCTTACCTGAACATCATGCGCTGCCACCGCATACA 420
Qy 141 GlySerTyLeuProValLeuValTPPheProGlyGlyValaPheTyThrGlySerAla 160
Db 421 GGCTCCAAAGCTCCCGCTTGGTGGTTCCTCCAGAGGTGGCTTCAAGATGCTGCTCAGCC 480
Qy 161 SerIlePheAspGlySerAlaLeuAlaAlaTyArgLysValLeuValValGln 180
Db 481 TCCATCTTTGAATGGGCTCCGCTGGCTGCTTGAAGAGCTGCTGGTGTGTGCTCCAG 540
Qy 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTPAAsnGlnHisAlaProGlyAsnTP 200
Db 541 TACCGGCTAGGAATTTGGTCTTTCACACATGGAGTACAGACATGCTCCGGGAACCTGG 600
Qy 201 AlaPheLysAspGlnValaAlaLeuSerTPValGlnLysAsnIleGluPhePheGly 220
Db 601 GCCTTCAAGGACCAAGTGGCTGCTGCTGGGTCCAGAAACATGAGATTCCTGGCT 660
Qy 221 GlyAspProSerSerValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSer 240
Db 661 GGGAGCCCAAGCTCTGTGACATCTTTGACAGTCCGGGAGCCCATGAAGTGTTCAGT 720
Qy 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGluSerGlyVal 260
Db 721 CTATATCTGTCTCCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGATGGGGTG 780
Qy 261 AlaIleIleProTyTPLeuGlnAlaHisAspTyArgLysSerGlyAspLeuGlnVal 280
Db 781 GCCATCATCCCTTATCTGGAGGCCCATGATTAAGAAAGTGAAGACCTGCAGAGTGTT 840
Qy 281 AlaHisPheCyGlyAsnAsnAlaSerAspSerGluAlaLeuAlaGlyGlyLeuArgThr 300
Db 841 GCACATTTCTGTGTAACAATGCTCAGACTCTGAAGCCCTGAGAGTCTGAGAGACA 900
Qy 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
Db 901 AAACCCCTCCAAAGAGCTGAGACCTCAGCCACAGAAACAAAGCTTTTCATCCAGTGCTT 960
Qy 321 AspGlyValaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysVala 340
Db 961 GATGGTCTTTCTTTCTTAATGAGCTCTGATGATTAATGCTCAAGAAAGATTTAAAGCA 1020
Qy 341 IleProSerIleIleGlyValaAsnAsnHisGlyCyGlyPheLeuLeuProMetLysGlu 360
Db 1021 ATTCCCTTCATCATCGAGTCAATTAACAAGATGGCTTCTGCTGCTTAAGAAAGAG 1080
Qy 361 AlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIle 380
Db 1081 GCTCTGAGATCTCTAGTGGCTCCAAACAGTCCCTTGGCTCATCTGATACAAACATC 1140

Qy 381 LeuHisIleProProGlnTyTPLeuHisLeuValaAlaAsnGlyTyPheHisAspLysHis 400
Db 1141 CTGCACATCCCGCTTCAGATTAATGACCTTGGGTGAATGAATCTTCACAGACAAAGCAC 1200
Qy 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyLysAspValPhePheValVal 420
Db 1201 TCCCTGACTGAATTCGAGACAGTCTTGGACCTTGTGAGATGTCTTTTGTGGTTC 1280
Qy 421 ProAlaLeuIleThrAlaArgTyHisArgAspAlaGlyValaProValTyPheTyGlu 440
Db 1261 CTTGACATGATCACAGCTCGATATCACAGATGCTGTGTGACCTGTCTACTTCTTAAG 1320
Qy 441 PheArgHisArgProGlnCyPheGlyLysAspThrLysProAlaPheValLysAlaAspHis 460
Db 1321 TTTGGGACCGGCTCAAGTCTTTGAGAACACAGAACCGGCTTTGTCAAGCCGACAC 1380
Qy 461 AlaAspGlyValaArgPheValPheGlyGlyValaPheLeuLysGlyLysPheIleValMetPhe 480
Db 1381 GCTGATGAAGTCCGCTTTGTGTGCTGGTGGTCTTCTGAAAGGGGACATTTGTAATGTT 1440
Qy 481 GluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetLeuTyTPValaThr 500
Db 1441 GAAGGACCAAGGAGAGAGAGATTAATCTGAGCCGAAAGTGAATGAATCTGGGCTAC 1500
Qy 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTPProAlaTyArgLys 520
Db 1501 TTTCTCGAAACCGGGAATCTTAATGGGACGACCTGTCTGTGGCCAGCTTAATCTG 1560
Qy 521 ThrGluGlnTyTPLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Db 1561 ACTGAGGACGATCTCCAGCTGAGCTTGAACATAGAGCTCGAGACAGATCAAAAGAACCG 1620
Qy 541 ArgValaAspPheTPThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
Db 1621 CGGGTGAATTTTGGACCAAGCAACATCCCTGATCTGTGCTCCACATGCTCCAC 1680
Qy 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCyVala 580
Db 1681 AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 581 Pro 581
Db 1741 CCT 1743

RESULT 2
US-10-023-515-1
Sequence 1, Application US/10023515
Patent No. 6664091
GENERAL INFORMATION:
APPLICANT: Curtiss, Roy A. J.
TITLE OF INVENTION: Siles-Santiago, Immaculada
TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
FILE REFERENCE: 10448-122001
CURRENT APPLICATION NUMBER: US/10/023,515
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,369
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/279,508
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2158
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (96) ... (1838)
US-10-023-515-1
Alignment Scores:

Pred. No.: 0 Length: 2158
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-10-023-515-2 (1-581) x US-10-023-515-1 (1-2158)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleuGlnPro 20
 DB 96 ATGCCACAGGAGCTTACTTCATCTGCTTCAACAATGCTGCTTTTCTTGATTCCTCAGGCC 155
 QY 21 LeuLeuGlnHlaArgGlnTrpGlyLeuThrGlyProSerAlaGlyGlyProGlnArg 40
 DB 156 CTGTTGGACACAGACAGTGGGAAAACCTGGGCTTCTGCTGAAAGGCCACAGAGAAC 215
 QY 41 ThrArgLeuGlyTrpIleGlnGlyLeuValThrValLeuGlySerProValProVal 60
 DB 216 ACCAGCTGGAGTGGATTCAGGGCAAGCAATGCTGCTGGAAAGCCCTGGCTGTG 275
 QY 61 AsnValPheLeuGlyValProPheAlaAlaProPheLeuGlySerLeuArgPheThr 80
 DB 276 AACGTTCTCTGGAGTCCCTTGTGCTCCCGCTGGAGTCCCTGGATTTACGAA 335
 QY 81 ProGlnProAlaSerProTrpAsnLeuArgGlyAlaThrSerTrpProAsnLeu 100
 DB 336 CCGACGCTGATCGCCCTGGAGTACCTTCCAGAAAGCCACCTCTTACCTTAATTTGTC 395
 QY 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHlaMetLeuValHlaSerProGly 120
 DB 396 CTCCAGAACTCAGAGTGGCTGCTTGAATCAACATGCTCAAGTGTATACCGGAA 455
 QY 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHlaSerPro 140
 DB 456 TTCGAGAGTCTGAGAACTGCTTCACTCAACATCTGAGCCCTGCCACCCGAACTCA 515
 QY 141 GlySerLeuLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySer 160
 DB 516 GCTTCAAGCTCCCGCTTGTGAGTTCCTCCAGAGGAGCTTCAAGATGCTCAAGCC 575
 QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpLeuAspValLeuValValGln 180
 DB 576 TCCATCTTGAAGGATCCGCTGCTGCTATGAGAGCTGCTGTTGTGTCCTCCAG 635
 QY 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHlaAlaProGlyAsnTrp 200
 DB 636 TACCGCTAGGAATTTGTTCTTCAACATGAGGATCAAGCATGCTCCGGGAACTGG 695
 QY 201 AlaPheLeuAspGlnValAlaAlaLeuSerTrpValGlnValAsnIleGluPheGly 220
 DB 696 GCTTCAAGAACCAAGTGGCTGCTGCTGCTCCAGAAAGCAATCAAGTTCTTCGCT 755
 QY 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSer 240
 DB 756 GGGAGCCCAAGCTCTGTACCATCTTTGGCGAGTCCGGGAGCCATAAGTCTTCTAGT 815
 QY 241 LeuIleLeuSerProMetAlaLeuGlyLeuPheHlaIleMetGlnSerGlyVal 260
 DB 816 CTTATACGTCTCCATGAGCCCAAGGCTTATCCAAAGCATCATGAGATGGAGTGG 875
 QY 261 AlaIleIleProTrpLeuGlnAlaHlaSerTrpGlyLeuSerGlnAspLeuGlnVal 280
 DB 876 GCATCATCTCTTACCTGAGGCCCATGATTAATGAGAGTGAAGACCTGCAAGTGGT 935
 QY 281 AlaHlaPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
 DB 936 GCACATTTCTGTGATCAATGCTGACATCTGAGGCCCTCTAAGGCTCTGAGAC 995
 QY 301 LysProSerLeuGlnLeuLeuThrLeuSerGlnTrpValSerPheThrArgValVal 320
 DB 996 AAACCTCAAGAGACTGCTGACCTCAGCCAGAAACAAAGTCTTCTCACTGAGTGGTT 1055

QY 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnIleValAlaPheVal 340
 DB 1056 GATGGTGGCTTTCTTCTTAATGAGCTCAGATCTATGTTCTCAGAAAGCATTTAAAGCA 1115
 QY 341 IleProSerIleIleGlyValAsnAsnHlaGlyCysGlyPheLeuLeuProMetLeuGly 360
 DB 1116 ATTCCTTCATCATGAGTCAATTAACACAGATGTGGCTTCTCTGCTCTTGAAGAG 1175
 QY 361 AlaProGlnIleLeuSerGlySerAsnIleSerLeuAlaLeuHlaLeuIleGlnAsn 380
 DB 1176 GCTCTGAGATCTCAATGAGCTCAACAGATCTTGGCTTCCATCTGATTAACAAACATC 1235
 QY 381 LeuHlaIleProProGlnTrpLeuHlaSerValAlaAsnGlyTrpPheHlaSerPhe 400
 DB 1236 CTGCACATCCCGCTCAGATTTGCACTTGGGTAAATGATATCTTCATCAACAGAC 1295
 QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
 DB 1296 TCCCTGACTGAATCCGAGACAGTCTTGGACTTGGAGATGTCTTGTGCTG 1355
 QY 421 ProAlaLeuIleThrAlaArgTrpHlaSerAspAlaGlyAlaProValTrpPheTrpGlu 440
 DB 1356 CTGCACTGATCAAGCTCGATATCAAGAGATGCTGGACCTGTCTACTTCTATGAG 1415
 QY 441 PheArgHlaArgProGlnCysPheGlnAspThrLysProAlaPheValIleValAsn 460
 DB 1416 TTTCCGACACCGGCTCAGTGTCTTGAAGACAGAGCCGGCTTTGTCAACCGACCC 1475
 QY 461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuValAspIleValMetPhe 480
 DB 1476 GCTGATGAAGTCCGCTTGTGTTGCGTGTGCTTCAAGGGGAGCACTGTATGTTTC 1535
 QY 481 GluGlyAlaThrGlnGlnGlyLeuLeuSerArgIleMetMetLeuTrpTrpAlaThr 500
 DB 1536 GAAGGAGCCACAGAGAGAGAGTACTGAGCCGAGAGAGATGAATACAGGCTAC 1595
 QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsn 520
 DB 1596 TTTGCTCAACCGGAAATCTTAATGGAAACAGCTGTCTGTGGCCAGCTTAATATCG 1655
 QY 521 ThrGlnIleTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuIleGluPro 540
 DB 1656 ACTGAGAGTACTCTCAGCTGGAATTGAACATGAGCTCCGACAGAGACTCAAAAGACG 1715
 QY 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHla 560
 DB 1716 CGGATGATTTTGGACAGCAACATCCCTGATCTGTGCTGCTCCGACATGCTCCAC 1775
 QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
 DB 1776 AGTCTCTTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
 QY 581 Pro 581
 DB 1836 CCT 1838

RESULT 3
 US-09-949-016-3799
 / Sequence 3799, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3799
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3799

Alignment Scores:
Pred. No.:      2,14e-136      Length:      2117
Score:          1230.00        Matches:      248
Percent Similarity: 63.84%      Conservative: 91
Best Local Similarity: 46.70%      Mismatches: 168
Query Match:     39.95%      Indels:      24
DB:              4            Gaps:         8

US-10-023-515-2 (1-581) x US-09-949-016-3799 (1-2117)

QY      31  GLYProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlySerGln 50
      132  GGGCAGGAGCTCAGCCAGTCCCATCCGGACACACACAGCGGAGGAGTGGGAGTCTT 191
QY      51  ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIa 70
      192  GTCCATGTAGAGGGGCGCCCAATGCCGGGGTCCAAACCTTCCTGGGAATTCATTGGCAAG 251
QY      71  ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 90
      252  CCACCTCTAGTCCGCTGGATTTGCAACCCCTGAGCCCTCGAATCTTGAGTGTG 311
QY      91  ArgGluAlaThrSerGlyProAsnLeuGlyLeuGlnAsn-----SerGlu 105
      312  AAGGATGAAACCAACCATCCGGCATGTGTCTACAGAGCACTCACCGCAGTGAAGTCAGAG 371
QY      106  TrpLeuLeuLeuAspGlnHisMetLeuValHisGlyProGlyPheGlyValSerGlu 125
      372  TTTCTT-----AGCCAGTTCAACATGACCTTCCTCCGATCTCATGTCTAG 419
QY      126  AspCysLeuTrpLeuAsnIleTyralAspAlaHisAlaAspThrGlySerLeuLeuPro 145
      420  GACTCCCTGTACTTACCTGACATCTACACGCGGCCCATAGCATGAAGGCTTAACTGCGG 479
QY      146  ValLeuValTrpPheProGlyValAlaPheLeuThrGlySerAlaSerIlePheAspGly 165
      480  GTGATGGTGTGGATCCAGCGTGTGCTGCTTTTGGCAGAGGCTTCTGTATGATGAT 539
QY      166  SerAlaLeuAlaAlaTyrglyAspValLeuValValValGlnTyrglyLeuGlyIle 185
      540  TCCATGCTGGCTGCTCTGGAGAACGTGGTGGTGCATCATCCAGTACCGCTGGGTGTC 599
QY      186  PheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLeuAspGln 205
      600  CTGGGCTTCTTCAGACATCGAGACAGACAGCACCGCAACCGGCACTGGGCTGAGCAAA 659
QY      206  ValAlaAlaLeuSerTrpValGlnLeuAsnIleGluPhePheGlyGlyAspProSerSer 225
      660  GTGGCTGACATCACTGGGCTGACAGACAGAAATATCCGCCACTTTGAGAGCAACCTGACCT 719
QY      226  ValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
      720  GTTACCAATTTTGGCGAGTCTGCGGGTGGACGAGTGTCTTGGCTTGTGTGTCCTCC 779
QY      246  MetAlaLeuGlyLeuPheHisIleValAlaIleMetGluSerGlyValAlaIleIleProTy 265
      780  ATATCCCAAGAGACTCTTCCACGAGCATATGAGAGTGGCGGCTCTGCGCCGCGC 839
QY      266  LeuGluAlaHisAspTyrglyLeuSerGluAspLeu-----GlnValValAlaHisAsp 283
      840  CTAATTGCTC-----AGCTCAGCTGATGTCTATCTCCACGGGTGGGCAACCTG 887
QY      284  CysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuArgCysLeuArgThrIleProSer 303
      888  TCTGCCGTGACCAAGATTGACTGTAGAGCCCTGTGGGCTGCTGCGGGCAAGAGTAA 947

```

```

QY      304  LysGluLeuLeuThrLeuSerGlnIleTyrsSerPheThrArgValValAspGlyAla 323
      948  GAGGAGATTCTTGCAATTAAACAGCTTTCAAGATGATCCCCGAGTGTGATGGGGTCTC 1007
QY      324  PhePheProAsnGluProLeuAspLeuLeuSerGlnValAlaPheValAlaIleProSer 343
      1008  TTCTGCCCCAGGACCCCCAGAGCTCTGGCTCTGCCACTTTCAGCTGTCCCTTAC 1067
QY      344  IleIleGlyValAsnAsnHisGlyCysGlyPheLeuLeuPro-----MetLys 359
      1068  ATTGTGGTGTCAACAAATGAAATTCGGCTGGCTCATCCCAAGTGCATGAGAGTCTAT 1127
QY      360  GluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsn 379
      1128  GATACCCAGAAAGAAATATGACAGAGAGGCTCCAGGCTGTCTGCAAGAAATTTTAAAG 1187
QY      380  IleLeuHisIleProProGlnTrpLeuHisLeuValAlaAsnGlyTrpPheHisAspLys 399
      1188  CTGCTGATGTGTGCTCTTCACTTTGTGACTGTGAGGAGGAGTACATTTGGAGCAAT 1247
QY      400  HisSerLeuThrGluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheVal 419
      1248  GGGGATCCCCAGACCTTCCAAAGCCAGTTCAGAGATGATAGGCGGACTCATGTTTGTG 1307
QY      420  ValProAlaLeuIleThrAlaArgTyrglyHisArgAspAlaGlyAlaProValTyrgly 439
      1308  ATCCCTGACTCCAAATGACATTTT---CAGTGTCCCGGGCCCTGTGTACTTAC 1364
QY      440  GluPheArgHisArgProGlnCysPheGluAspThrValPheValAlaPheValAlaAsp 459
      1365  GAGTTTCAGATCAAGCCCACTGCTGACAAATCAAGGACGACGACCAATGAAGGACAG 1424
QY      460  HisAlaAspGluValArgPheValPheGlyValAlaPheLeuLysGlyAspIleValMet 479
      1425  CATGTGATAGGCTTCTCTTGTGTTTC---AGAATTTCTTGGGGCAACTGAACTAA 1481
QY      480  PheGluGlyAlaThrGluGluGluValLeuLeuSerArgLysMetLeuTyrglyTrpAla 499
      1482  TTC-----ACTAGGAAGAGAGACAGCTTACGAGAAAGTGAAGTGAAGTACGGGCG 1532
QY      500  ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyraen 519
      1533  AACTTGGCGAAGATGGAACCCCAATGGCGAGGGTCTCCACATGGCGGCTGTTCAGC 1592
QY      520  LeuThrGluGluTyrgluGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlu 539
      1593  CAGAGAGAGCAATCACTGCACTGAACCTTACAGCTGCGGTGGCGGCTCTGAAGGCC 1652
QY      540  ProArgValAspPheTrpThrSerThrIlePro 550
      1653  CACAGGCTCAAGTTCTGGAAGAAAGGCGTGGCC 1685

RESULT 4
US-09-949-016-555
; Sequence 555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 555

```

```

; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-555

Alignment Scores:
Pred. No.: 2,23e-136 Length: 2169
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 39.95% Indels: 24
DB: 4 Gaps: 8

US-10-023-515-2 (1-581) x US-09-949-016-555 (1-2169)
QY 31 GYProSerAlaGluGlyProGlnArgAntThrArgLeuGlyTyrP1leGlnGlyArgGln 50
DB 132 GGCACAGACTGACCCAGTCCATCCGGACACACACGCGGACAGGTGCTGGGAGTCTT 191
QY 51 ValThrValLeuGlySerProValProValAanValPheLeuGlyValPProPheAla 70
DB 192 GTTCATGTGAAGGGGCCCAATCCGGGGTCCAACTTCTTGGAAATTCATTGCGCAAG 251
QY 71 ProProLeuGlySerLeuArgPheThrAanProGlnProAlaSerProTTrpAspAanLeu 90
DB 252 CCACCTTAGGTCCGCTGGATTGTGACCCCGTGAAGCCCTGGAATCTTGAGTGTGTG 311
QY 91 ArgGlnAlaThrSerTyrProAanLeuGlyLeuGlnAan-----SerGlu 105
DB 312 AGGATGGAACCAACCATCCGGCATGTGTCTACAGAGCATCCGACGAGTGAGTCAGAG 371
QY 106 TrpLeuLeuLeuAanPheGlnHisMetLeuGlyValHisTyrProLysPheGlyValSerGlu 125
DB 372 TTTCCTT-----AGCCAGTTCAACATGACCTTCTCCGATCCCATGCTCAG 419
QY 126 AspGlyLeuTyrLeuAanL1eTyrAlaProAlaHisAlaAspThrGlySerLysLeuPro 145
DB 420 GACTCCCTGTACTCTGACATCTACACGCGGCCCATAGCATGAAGGCTCTAACTGCGCG 479
QY 146 ValLeuValTTrpPheProGlyGlyValAphelyrThrGlySerAlaSerL1ePheAphGly 165
DB 480 GTGAAGGTGTGTGATCCACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 539
QY 166 SerAlaLeuAlaAlaTyrGlnAspValLeuValValValGlnTyrArgLeuGlyLe 185
DB 540 TTCATGCTGGCTGCTTGTGAGAACG1GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
QY 186 PheGlyPhePheThrThrTTrpAspGlnHisAlaProGlyAsnTTrpAlaPheLysAspGln 205
DB 600 CTGGGCTTCTTACGACATGAGACACGACACGACGCGAACTGGGGCTTACCTGAGACCA 659
QY 206 ValAlaAlaLeuSerTyrValGlnLysAanL1eGluPhePheGlyGlyAspProSerSer 225
DB 660 GTGGCTGACATCGCTGGGTCCAGCAAGATATCGCCACTTGTGAGGCAACCTGACCGT 719
QY 226 ValTTrpL1ePheGlyGlySerAlaGlyAlaL1eSerValSerSerLeuLeuLeuSerPro 245
DB 720 GTACACCAATTTTGGGAGCTGGGAGTGGGACGAGGTGTCTTCCCTGTGTGTGTGTGTGTGT 779
QY 246 MetAlaLysGlyLeuPheHisLysAlaL1eMetGlySerGlyValAlaL1eLeuProTyr 265
DB 780 ATATCCCAAGACCTTCCACGAGACCATCATGAGAGTGGGTGGCTCTCTGCGCGCG 839
QY 266 LeuGlnAlaHisAspTyrGlnLysSerGluAanLeu-----GlnValValAlaHisPhe 283
DB 840 CTCATATGCC-----AGCTCAGCTGATGTATCTTCCACCGTGTGTGTGTGTGTGTGTGT 887
QY 284 CysGlyAanAanAlaSerAspSerGlnAlaLeuLeuAanGlyLeuArgThrLysProSer 303
DB 888 TCTGCTGTGACCAAGTTGACTCTGAGGCTGTGGGCTGTGCTGTGCGGGCAAGAGTAA 947
QY 304 LysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValValAspGlyAla 323

```

```

DB 948 GAGAGATTCTTTCATTAAACAAGCTTTCAAGATGATCCCGGAGTGTGATGGGCTC 1007
QY 324 PhePheProAanGlnProLeuAspLeuSerGlnLysAlaPheLysAlaL1eProSer 343
DB 1008 TTCCTGCCCCAGGACCCCCAGAGCTGTGGCTCTTCCCGACTTTCAGCTGTCCCTTAC 1067
QY 344 L1eL1eGlyValAanAanHisGlyCysGlyPheLeuLeuPro-----MetLys 359
DB 1068 ATTGTGTGTGTCACACAAATGAAATTCGCTGTGCTCATCCCAAGTCTATGAGATCTAT 1127
QY 360 GlnAlaProGlnL1eLeuSerGlySerAanLysSerLeuAlaLeuHisL1eGlnAan 379
DB 1128 GATACCCAGAAAGAAATGACAGAGAGGCTTCCAGGCTCTCTCAGAAATGTTAACG 1187
QY 380 L1eLeuHisL1eLeuProGlnTyrL1eHisL1eLeuValAlaAanGlnTyrPheHisAspLys 399
DB 1188 CTGTGATGTGTCTCTTCAATTTGTGTGATCTGTGTGAGGAGAGATCACTTGGGACAAAT 1247
QY 400 HisSerLeuThrGlnL1eAspSerLeuLeuAanLeuGlyAspValPhePheVal 419
DB 1248 GGGGATCCCCAGACCTCCAAAGCGAGTTCACAGAGATGATGGCGGACTCCATGTTTGTG 1307
QY 420 ValProAlaLeuL1eThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
DB 1308 ATCCCTGCACTTCAAGTACATATTT---CAGTGTCCCGGGCCCTGTGTATCTTCTAC 1364
QY 440 GlnPheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAsp 459
DB 1365 GAGTTCACAGATCAGACCCAGCTGGCTCAAGAACATCAGGCCCGCACATGAAGCAGAC 1424
QY 460 HisAlaAspGlnValArgPheValPheGlyGlyAlaPheLysGlyAspL1eValMet 479
DB 1425 CATGTGTGATGACTTCTTGTGTGTTC---AGAACTTCTTGTGGGGCAACTACATTA 1481
QY 480 PheGlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTyrTrpAla 499
DB 1482 TTC-----ACTGAGAAAGAGAGACAGCTTAACAGAGATGATGAAATGATCGGGCC 1532
QY 500 ThrPheAlaArgThrGlyAsnProAanGlnLysAanPheLeuSerLeuTTrpAlaTyrAan 519
DB 1533 AACTTGGAGAAATGAGAACCCCAATGCGAGGGCTGCCACACTGGCGCTGTGTGAC 1592
QY 520 LeuThrGlnGlnTyrLeuGlnLysLeuAanMetSerLeuGlyGlnArgLeuLysGln 539
DB 1593 CAGAGAGAGCAATCTCTGACTGAACCTTACAGCTGTGGGTGGGCTGTGAAGGCC 1652
QY 540 ProArgValAspPheTTrpThrSerThrL1ePro 550
DB 1653 CACAGGCTCCAGTTCTTGAGAAAGCGCTGCC 1685

RESULT 5
US-09-595-682B-27
; Sequence 27, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danke, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2191
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-595-682B-27

Alignment Scores:

Pred. No.:	2,27e-136	Length:	2191
Score:	1230.00	Matches:	248
Percent Similarity:	63.84%	Conservative:	91
Best Local Similarity:	46.70%	Mismatches:	168
Query Match:	39.95%	Indels:	24
DB:	4	Gaps:	8

US-10-023-515-2 (1-581) x US-09-595-682B-27 (1-2191)

```
QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPrlLeuGlnGlySerGln 50
Db 154 GGCACGAGATCAGCAGCAGTCCCATCCGACACACACGCGGAGAGGTCTGGGGAATCTT 213
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 70
Db 214 GTCCATGTGAAGGCGCAATGCCGGGGTCCAAACCTTCCTGGGAATTCATTGGCCAAAG 273
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 90
Db 274 CCACCTTAGTCCGCTGCGATTGGACATCCCTGAGCCCTCCGATCTTGAGAGTGTG 333
QY 91 ArgGluAlaThrSerTyrProAsnLeuGlyLeuGlnAsn-----SerGln 105
Db 334 AGGAGTGAACCAACCATCCGCGCATGTGCTACAGAGCACTCACCGCAGTGAAGTCAAG 393
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuValHisTyrProValPheGlyValSerGln 125
Db 394 TTTCCTT-----AGCCAGTTCAACATGACCTTCTCCGACTCCATGCTCAAG 441
QY 126 AspGlyLeuTyrLeuAsnLLeTyrAlaProAlaHisAlaAspThrGlySerLeuPro 145
Db 442 GACTGCTGTACTCAGCATCTACACGCGCGCCATAGCAGTGAAGGCTTAACTGCG 501
QY 146 ValLeuValThrPheProGlyValAlaPheLeuThrGlySerAlaSerLeuPheAspGly 165
Db 502 GTGATGGTGTGATCCACGCTGTGCTGCTTTGTTTGGCATGGCTTCTTGATATATG 561
QY 166 SerAlaLeuAlaAlaTyrGlnAspValLeuValValAlaGlnTyrArgLeuGlyLe 185
Db 562 TCCATGCTGCTGCTGCTGAGAACGCTGGTGTGCTATCATCTACCTGAGTGTG 621
QY 186 PheGlyPhePheThrThrTTrpAspGlnHisAlaProGlyAsnTrpAlaPheLeuAspGln 205
Db 622 CTGGGCTTCTTCAAGCATGAGACACAGACACGCAACCGGCAACTGGGCTTACTGACCAA 681
QY 206 ValAlaAlaLeuSerTTrpValGlnLysAsnLLeGluPheGlyValAspProSerSer 225
Db 682 GTGGCTGACATACGCTGGTCCAGCAGAAATTCGCCCATTTGAGAGGCAACCTGACCGT 741
QY 226 ValThrLeuPheGlyLeuSerAlaGlyAlaLLeSerValSerSerLeuLeuSerPro 245
Db 742 GTCAACATTTTGGCGAGTCTGCGGGTGGCACAGTGTGCTTGCCTTGTGTGTCCTCC 801
QY 246 MetAlaLysGlyLeuPheHisAlaAlaLLeMetGlnSerGlyValAlaAlaLLeProTyr 265
Db 802 ATATCCCAAGACTCTTCCACGAGCCATCTGAGAGTGGCGTGGCCCTCCGCGCGC 861
QY 266 LeuGluAlaHisAspTyrGlnLysSerGlnAspLeu-----GlnValAlaHisPhe 283
Db 862 CTCAATGCC-----AGCTCAGCTGATCTCATCTCCACGGTGGTGGCAACCTG 909
QY 284 CysGlyLysAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThrLysProSer 303
Db 910 TCTGCTGTGACCAAGTGTGACTGTGAGGCTTGTGGCGCTGCGCGGCAAGAGTAA 969
QY 304 LysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValValAspGlyAla 323
Db 970 GAGAGATTTTTCGAATTAACAAGCTTTCAAGATATCCCGAGAGTGGTGGAGTGGTGC 1029
QY 324 PhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAlaLLeProSer 343
```

```
Db 1030 TTCCTGCCAGGACCCCAAGAGCTGCTGGCTTCCGCACTTTCAGCTGTCCCTAGC 1089
QY 344 IleIleGlyValAlaAsnHisGluCysGlyPheLeuLeuPro-----MetLys 359
Db 1090 ATTGTGGTGTCAACAACTAATATTTGGCTGCTCATCCCAAGTCAATGAGATCTAT 1149
QY 360 GluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLLeuIleGlnAsn 379
Db 1150 GATACCCAGAAAGAAATGAGACAGAGGCTCCAGCGTGTCTTCGAGAAATGTTTAAC 1209
QY 380 IleLeuHisLLeProProGlnTyrLeuHisLLeuValAlaAsnGlyTyrPheHisAspLys 399
Db 1210 CTGCTGATGTGGCTCTTCACTTGTGACTCTGAGGAGAGTACATTGGGAGCAAT 1269
QY 400 HisSerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
Db 1270 GGGGATCCCAAGACCTTCCAAAGCCAGTTCCAGAGATGATGGCGGACTCCATGTGTG 1329
QY 420 ValProAlaLeuLLeThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
Db 1330 ATCCCTGACTCCAAATGACATATTT---CAGTGTCCCGGGCCCTGTGTACTTCTAC 1386
QY 440 GluPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAsp 459
Db 1387 GAGTTTCAAGATCAGCCAGCTGCTGACAGAACATCTGACCGCAGCATGAAGGACAG 1446
QY 460 HisAlaAspGluValaArgPheValPheGlyAlaPheLeuLysGlyAspLLeuLLeMet 479
Db 1447 CATGTGATGAGCTTCTTTGTTTTC---AGAAATTTCTTGGGGGCAACTATTA 1503
QY 480 PheGluGlyAlaThrGluGluGluLysLLeuLeuSerGlyLysMetLysTyrTrpAla 499
Db 1504 TTC-----ACTAGAGAAAGAGAGCAGCTTAAAGAGAAATGATGAACTAGTGGCC 1554
QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuThrProAlaTyrAsn 519
Db 1555 AACTTTCGAGAAATGGAAGAACCCCAATGCGAGGCTGTCCACACTGCGCCCTGTTCGAC 1614
QY 520 LeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGln 539
Db 1615 CAGAGAGAGCAATATCTGACGTGAACCTACAGCTGCGTGGCGCGGCTCTGAAGGCC 1674
QY 540 ProArgValaAspPheThrThrSerThrLLePro 550
Db 1675 CACAGGCTCAGTTTCTGAAAGAGCGCTGCC 1707

RESULT 6
US-09-595-682B-20
/ Sequence 20, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danke, Mary K.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR FILING DATE: 2000-01-16
/ PRIOR APPLICATION NUMBER: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR APPLICATION NUMBER: PCT/US99/03171
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 20
/ LENGTH: 1717
/ TYPE: DNA
/ ORGANISM: Oryctolagus cuniculus
US-09-595-682B-20
Alignment Scores:
```

Pred. No.:	7,83e-126	Length:	171
Score:	1140.50	Matches:	251
Percent Similarity:	56.92%	Conservative:	78
Best Local Similarity:	43.43%	Mismatches:	182
Query Match:	37.04%	Indels:	67
DB:	4	Gaps:	13

US-10-023-515-2 (1-581) X US-09-595-682B-20 (1-1717)

[illegible]

```

Oy 313 -----ThrylsSerPheThrArgValaIAsrGly 322
Db 923 CTAAATCTAGTGGCCGACCCCAAGAGAACACCCCTTCTGCACTAGTGAATGG 982
Oy 323 AlaPheProAengIuProLeuAsrPheLeuSerGlnIuValaPheValaIlePro 342
Db 983 GTGCTGCTGCAAAAGACCTGGACGATGTTCTGGCAGAGAAAGAAATCAACATGCTGCC 1042
Oy 343 SerIleIleGlyValaIAsnAsnIleGluCysGlyPheLeuLeuProMet----- 358
Db 1043 TACATGGTGGAAATCAACACAGCAGAGATTTGGCTGATTTCCCAATGCCAATGCTGGCC 1102
Oy 359 -----LysGluIAsrProGluIleLeuSerGlySer 368
Db 1103 TATGCACTCTTGAGGCGAAACCTGGACCGAGAGACAGCTACAGAACTCTTGGAAGTCC 1162
Oy 369 AsnLysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProGlnIuTrLeu 388
Db 1163 -----TACCCTATTGTCAATGCTCTAGAGACTGACT 1195
Oy 389 HisLeuValaIAsnGluTrPheHisAsrPheHisSerLeuThrGluIleAsrSer 408
Db 1196 CCAATGGCCCACTGAGATGTAATTGTAGAGGAGCAGATACCCCTGCAAAAGAAAGACTTG 1255
Oy 409 LeuLeuAsrPheLeuGluIAsrValaPhePheValaIProAlaLeuIleThrAlaIAsrTr 428
Db 1256 TTCTGTGACATGCTGTGACAGATTGTGTAATTGGTGTCCATCTGGAATGTGGCTGTGCAC 1315
Oy 429 HisArgAsrAlaGlyAlaIAsrProValaTrPheTrGluPheArgHisAsrProGlnCysPhe 448
Db 1316 CACGAGATGCTGGAGCCCAACCTATATGTATAGATATCGGATATCCGCCAAGTTCTTCA 1375
Oy 449 GluAsrThrTrpProAlaPheValaIAsrPheHisAlaAsrGluValaIAsrPheValaPhe 468
Db 1376 TCAGACATGAGACCCCAAGACAGTATAGGGAGACCATGAGATGAGATCTTCTGTCTTA 1435
Oy 469 GlyIAsrAlaPheLeuLysGlyIAsrPheIleValaMetPheGluGlyAlaThrGluGluLys 488
Db 1436 GGACCCCGTTTTTAAAA-----GAGGTGTCACAGAGAGAAATC 1477
Oy 489 LeuLeuSerArgLysMetLeuLysTrpTrpAlaThrPheAlaArgThrGlyAsnProAsn 508
Db 1478 AAATGTAGCAAGATGTGATGAAATACTGGGCCAATCTTGTAAGATGGGAATCCCAAT 1537
Oy 509 GlyAsnAsrPheLeuSerLeuTrpProAlaTrpAsnLeuThrGluGlnIuTrpLeuGlnLeuAsr 528
Db 1538 GGAAGAGGCGCTTCTCTCAATGGCCAGCATATGACTACAGAGAAAGGTATCTGCACATTTGA 1597
Oy 529 LeuAsnMetSerLeuGlyIAsnArgLeuLysGluProArgValaAsrPheTrpThr 546
Db 1598 GCCACCAACCAAGCAGGCCCAAGAAATGAAAGACAGAGAGTGGCTTCTTGACT 1651

RESULT 7
US-09-264-737-3
; Sequence 3, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; FILE REFERENCE: 38-21(10551) RE3 Pyridine Enzymes
; CURRENT APPLICATION NUMBER: US/09/264,737A
; EARLIER FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Rabbit
; US-09-264-737-3

```


Alignment Scores:

Pred. No.:	1,34e-125	Length:	1701
Score:	1138.50	Matches:	250
Percent Similarity:	56.62%	Conservative:	75
Best Local Similarity:	43.55%	Mismatches:	176
Query Match:	36.98%	Indels:	73
DB:	3	Gaps:	14

US-10-023-515-2 (1-581) x US-09-264-737-3 (1-1701)

```

QY      12  TTPCyArPhePheLeuIleuGlnProLeuLeuGlnIleArGlnITPGLyLysThrGly 31
DB      25  TGGGCTCTTC-----CTTGCAGCCCTGCACCCGATGGGTAC-----60

QY      32  ProSerIaGluGlyProGlnIaArgAnthArGleuGlyTrPrlleGnIyLysGlnVal 51
DB      61  CCCCTCCGACCA---CCTGTGTTGACACTGTAAAGGGAAAGTCTGGGGAAAGTTCGTC 117

QY      52  ThrValIleuGlySerProValProValAsnValPheLeuGlyValProPheIaIaPro 71
DB      118  AGCTGAAGAGATTGGACAGCCGTCGCGCTTCTTGGAGATCCCTTCGCAAGCCC 177

QY      72  ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProITrapAsnLeuArg 91
DB      178  CCTCTGGATCCCTGAGGTTTGACACCAAGCTGCAGAAATCATGAGACCGTGAAG 237

QY      92  GluAlaThrSerITrapAsnLeuCyLeuGlnAsnSerGluITrLeuLeuLeuArgIn 111
DB      238  AACACACACCTCCACCTCCCATGTCTCCCAAGAGCA-----GTATCAGGG 285

QY      112  HlaMeLeu-----LysValIleArTyPro---LysPheGlyVal 123
DB      286  CATATGCTCTCGAGCTTTCACCAAGAAAGAAACATCCCTTAAGTT----- 339

QY      124  SerGluArPheLeuITrPheLeuAsnIleITrAlaProAlaIleAsnThrGlySerTy 143
DB      340  TCTGAAGATGCTTCTTACCTGAAATTTTACACCCCTGCTGACCTGACAAAGAGAGCGG 399

QY      144  LeuProValIleuValITrPheProGlyGlyAlaPheLysITrGlySerAlaSerITrPhe 163
DB      400  CTGCGCGGTGATGTTGATGCATGAGAGGTGTCATGATGGTGGTGGAGCATCAACCTAT 459

QY      164  AspGlySerAlaLeuAlaLeuITrGlyLysArValIleuValValGlnITrArgLeu 183
DB      460  GATGACCTGGCTCTTCTGCCCCATGAAACGTGGTGGTGAACCATTCAGTACCGCTG 519

QY      184  GlyITrPheGlyPhePheThrITrPArGlnIleAlaProGlyAsnITrAlaPheLys 203
DB      520  GGCATCTGGGATTTCTTACGACACAGAGATGAGACACCGGAACTGGGGTCACTTG 579

QY      204  AspGlnValAlaAlaLeuSerITrValGlnLysAsnIleGluPheGlyGlyAsnPro 223
DB      580  GACCAAGGTGCTGGCTGGCTGGTGCAGACAAATATTCCAATTTGGAGGGAGCA 639

QY      224  SerSerValITrITrPheGlyGlySerAlaGlyAlaIleSerValIleSerLeuIleLeu 243
DB      640  GGCCTGTGACATCTTTGAGAGTCAACAGAGAGGTCAAAAGTGTCTATCTTCTATTA 699

QY      244  SerITrMetAlaLysGlyLeuPheIleLysAlaIleMetGlySerGlyValAlaIleIle 263
DB      700  TCCCCCTGTGACCAAGATCTCTTCATCGAGCAATTTCCGAGATGGCGTGGCCCTCT 759

QY      264  ProITrLeuGlnIleIleArGlyGlyLysSerGlyLysPhe-----LeuGlnValAla 281
DB      760  TCCAGTCTC-----TTACGAAAGAACCAAGTCTTGGTGAAGAAATTTGCC 807

QY      282  HisPheCyGlyAsnAsnAlaSerAspSerGlyAlaLeuLeuArgCyLeuArGTrITrLys 301
DB      808  ATCGAAGCTGGGTGTAACACCACTCGCTGTATGTTCACTGCTCGGCCCAAGAG 867

QY      302  ProSerITrGlyLeuLeu-----ThrLeuSerGlnLys-----312

```

```

DB      868  ACAGAGAGAAATCATGAGGTGACATTGAATGAATTAATTAATGCTCTAGATCTACTT 927
QY      313  -----ThrLysSerPheITrArgValIleAspGlyAlaPhePhePro 326
DB      928  GGGACCCCAAGAAAGAACCCCTTCTGTGACCTGTGATTTGATGGGTGCTGTGCA 987
QY      327  AsnGluProLeuAspLeuLeuSerGlnITrLysAlaPheLysValIleProSerITrleGly 346
DB      988  AAAGCACTCGAGAGATTTCTGSCAGAGAAATKCAACATGCTGCCCTACATGGTGGGA 1047
QY      347  ValAsnAsnIleGlyCyGlyPheLeuLeuProMet-----358
DB      1048  ATCAACACAGAAAGATTTGGCTGATATCCCAATGCAAAATGCTGGGTATCACTCTCT 1107
QY      359  -----LysGluAlaProGluIleLeuSerGlySerAsnLysSerLeu 372
DB      1108  GAAGGCAAACTGCACAGAAACAGTACAGAACTCTTGTGAAGTCC-----1155
QY      373  AlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnITrLysHisLeuValAla 392
DB      1156  -----TACCCATTTGCAATGCTCTTAAGAGAGTGACTCCAGTGGCCACT 1200
QY      393  AsnGluITrPheIleArLysITrHisSerLeuITrGluIleArgAspSerLeuLeuArg 412
DB      1201  GAGAAATTTTGAAGAGGACAGATGACCTGTGCAAAAGAAAGACTTTGTTCTTGACATG 1260
QY      413  LeuGlyAspValPhePheValIleProAlaLeuIleITrAlaArgITrHisArGAspAla 432
DB      1261  CTTGCAGATTTGTTATTTGGTGTCCATCTGTGAATGTGGCTGTGCACCAAGATGCT 1320
QY      433  GlyAlaProValITrPheITrGlyLysPheArGHisITrArgProGlnCyPheGluArITrLys 452
DB      1321  GAGCCCCCACTATATGTTAGTATCGGTATCGCCCAAGCTTCTCATCAGACATGAGA 1380
QY      453  ProAlaPheValIleValIleAsnHisAlaAspGluValITrArgPheValITrPheGlyAlaPhe 472
DB      1381  CCCAAGACAGTGTATGAGGACCATGAGATGAGATCTTCTGTCTTAAGAGCCCCGTTT 1440
QY      473  LeuLysGlyAspIleValIleMetPheGluGlyAlaITrGluGlnGlyLysLeuLeuSerArg 492
DB      1441  TTAAA-----GAGGGTGCACAGAAAGAGATCAAACTGAGCAAG 1482
QY      493  LysMetMetLysITrITrAlaITrPheAlaArgITrGlyAsnProAsnGlyLysAspLeu 512
DB      1483  ATGTGATGAATTAATCTGGGCCAACTTGTCTAGAAATGGAAATCCAAATGAGAGGGCTT 1542
QY      513  SerLeuITrProAlaITrAsnLeuITrGluGlnITrLysGlnLeuAspLeuAsnMetSer 532
DB      1543  CTTCAATGGCCAGCATATGATCAAGAAAGTTACTTGCAGATTTGAGGCCACACCCAG 1602
QY      533  LeuGlyITrArgLeuLysGluITrArgValIleAspITrITrThr 546
DB      1603  GCAGCCCAAGAACTGAAGAACAAAGAAAGTGGCTTCTGGAAT 1644

RESULT 8
US-09-799-451-155
Sequence 155, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyun
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungqing
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhimei
APPLICANT: Wang, Dunrui

```


APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 155
LENGTH: 965
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (128) .. (925)
US-09-799-451-155

Alignment Scores:
Pred. No.: 5,85e-125 Length: 965
Score: 1129.50 Matches: 222
Percent Similarity: 80.14% Conservative: 0
Best Local Similarity: 80.14% Mismatches: 4
Query Match: 36.68% Indels: 52
DB: 4 Gaps: 1

US-10-023-515-2 (1-581) x US-09-799-451-155 (1-965)

QY 31 GlyProSerAlaGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLeuGln 50
DB 287 GGGCTTCTGCTGAGAGGGCCAGAGAGACACAGGCTGGATTCAGGCGCAGACGA 346
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 347 GTCACTGCTGCTGGAGAGCCCTGCTGCTGGAACGTTCCTCGAGAGCCCTTTCCTGCT 406
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeu 90
DB 407 CCCCCGCTGGAGTCCCTCGATTACGAACCCGACCCGATCCGCTCGGATTAACCTG 466
QY 91 ArgGlnAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGlyTrpLeuLeuAsp 110
DB 467 CGAGAGACCACTCTTACCTTAATTTGCTCCAGAACTCAGATGGCTGCTCTTAAGT 526
QY 111 GlnHisMetLeuValValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeu 130
DB 527 CAACGACATGCTCAAGGTGATTACCCCAAAATTCGAGGTGTCAAGAGACTGCTTAACCTG 586
QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhe 150
DB 587 AACATCTATGCGCTGCCCAAGCCGATACAGGCTCAAGCTCCCGCTTGGTGTGCTTC 646
QY 151 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa 170
DB 647 CAGAGAGGCTCTCAAGACTGCTCAGCTCCTCATCTTGTAGTGGTCCGCTCGCTGCTG 706
QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
DB 707 TATGAGAGCGTGTGCTGTGTGCTCCAGTACCGGCTAGGAATTTGCTTCTTCA 766
QY 191 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer 210
DB 767 TA----- 768
QY 211 TrpValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
DB 768 ----- 768
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 250
DB 769 -----CTGCTCCCATGCGCAAGGCTTA 792

QY 251 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 270
DB 793 TTCCACAAAGCCATCATGAGAGTGGGTGGCCATCTTCACTGAGGCCCATGAT 852
QY 271 TyrGluLysSerGluAspLeuGlnValAlaHisPheCysGlyAsnAsnAlaSerAsp 290
DB 853 TATGAGAGCGTGGAGAGCTGTCAGAGCGGTGCAATTTCTGTGTAAACAATCGTCAAG 912
QY 291 SerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeu 307
DB 913 TCTGAGCGCTGCTGAGGTGCTGAGGACAAACCTCCAGAGACTGCTG 963

RESULT 9
US-09-949-016-2515
Sequence 2515; Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2515
LENGTH: 2802
TYPE: DNA
ORGANISM: Human
US-09-949-016-2515

Alignment Scores:
Pred. No.: 8,25e-84 Length: 2802
Score: 793.00 Matches: 198
Percent Similarity: 48.00% Conservative: 114
Best Local Similarity: 30.46% Mismatches: 226
Query Match: 25.76% Indels: 112
DB: 4 Gaps: 19

US-10-023-515-2 (1-581) x US-09-949-016-2515 (1-2802)

QY 12 TrpCysPhePheLeuIleLeuGlnProLeuLeuGly-----HisArg 25
DB 256 TGTGCTTGGAGACACTGCTCTCTACATTTGCTGGGACCTGATGCCCTGCCGGAAC 315
QY 26 GlnTrpGlyLysThrGlyPro----- 32
DB 316 ATGTGGCTGGCGCTTGGCGCGCGCTGCTGCTGAGGCCCAAGCCACGCTGGCAGG 375
QY 33 -----SerAlaGlu 35
DB 376 AGCTTGCTCTCAACCTGTGTTCTCACTTGGCGCTGAGGCGCAGTACCAAGGCCCA 435
QY 36 GlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLysGlnVal----- 51
DB 436 GCACCCACAGTCAACACTCACTTGGAGAGCTTGAAGGTTGCCCGAGTACCACTGCCACT 495
QY 52 ThrValLeuGlySerProValAsnValPheLeuGlyValProPheAlaIaPro 71
DB 496 GAGATCTGGGG-----CTGTGACCAAAATACCTGGGGGTGCTCCATGACACTCTCC 546
QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 91
DB 547 CCGATCGCGCGAGAAAGTTCCTGCTGCTCCCTGAAACCAACCCCATCTGTGCGGCGATCCGG 606
QY 92 GlnAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGlu----- 105

```

Db      607 AACCCACACACATTCCCGGAGTGTGCCCCAGAAACATCCACAGCTGTGCCGAAGTC 666
Qy      106 -----TrrleuLeuAapGlnHisMetLeuValHisTyrProLysPhe 121
Db      667 ATGTGCGCGGCTGTTCATCTCCAACTGGATTCCTGCTACTCTTCATCCAGAG--- 723
Qy      122 GlyValSerGluAapCysLeuTyrLeuAapHisTyrAlaProAlaHisAlaAapThrGly 141
Db      724 -----CCCAACGAAAGACTGTCTTACTGAAAGCTCATATGTGCCAGCGAGATGATCCGGC 780
Qy      142 SerLys----- 143
Db      781 GCTAAGAAACAGCGCGAGACTTAGCCGATATATGACCGGGATGAAATGAAACATCCCG 840
Qy      144 -----LeuProValLeuValTrrPheProGlyGlyValAapLeuSrrGlySer 159
Db      841 GACAGTGTGTCTAAACCCCTCATGTGTCTATCATCCAGAGGCTCTTACATGAAAGGAGAA 900
Qy      160 AlaSerLeuPheAapGlySerAlaLeuAlaAlaTyrGluAapValLeuValValVal 179
Db      901 GGCACATGATGATGAGCAGCATCTCGCCAGTTATGGCAATGTCATCTCATACCCCTC 960
Qy      180 GlnTyrArgLeuGlyLysPheGlyPhePheThrThrTrrAapGlnHisAlaAapGlyAan 199
Db      961 AACTATCGGCTTGAAGTGTAGGTCTTCTGAGTACTGAGATCAAGGCTCCAAAGGCAAC 1020
Qy      200 TrrAlaPheLysAapGlnValAlaAlaLeuSerTrrValGlnLysAapHisLeuPhePhe 219
Db      1021 TATGGGCTCTTGACAGATCCAGAGCCCTCCGCTGGGTAGAGAAATATTCCTTTTC 1080
Qy      220 GlyGlyAapProSerSerValThrLysPheGlyLysSerAlaGlyAlaLysSerValSer 239
Db      1081 GGGGGAGACCCCGCGGATCACTGTCTTGGCTCGGCAATGGTGTCATCCGCTCAGC 1140
Qy      240 SerLeuLysLeuSerProMetAlaLysGlyLeuPheHisValAlaLysLeuGlySerGly 259
Db      1141 CTCCTCAAGTGTTCATCATCATCAAGAGGACTTTTCCAGAGGCCATCATCAAGAGTGC 1200
Qy      260 ValAlaLysLeuProTyrLeuGlnAlaHisAapTyrGluLysSerGluAapLeuGlnVal 279
Db      1201 TCTGTCTGTCCAGCTGG-----GCTGTGACTCAACCAAGCTGAGATGACACCGCTG 1254
Qy      280 ValAlaHisPheCysGlyAanAapHisAapSerSerGlyAlaLeuLeuAapCysLeuAap 299
Db      1255 CTGGAGAACAAAGTGGCTGTATGTCTGTGACACCGGATATGTGTGACTGTCTTCGG 1314
Qy      300 ThrLysProSerLysGlyLeu-----LeuThrLysSerGlnLysThrLysSer 315
Db      1315 CAAGAAGTGCACAGAGCTGGTAGAGCAGACATCCAGCCAGCCCGCTACACAGTGGCC 1374
Qy      316 PheThrArgValAlaAapGlyAlaAapPheProAapGlnProLeuAapLeuSerGln 335
Db      1375 TTTGGCCCTGTGATGTATGATGTATGTCTATCTCTGACCTTGAGATCTTCATGAGACAG 1434
Qy      336 LysAlaAapLeuValAlaLysSerLysLeuGlyValAanAapHisGlu---CysGlyPhe 354
Db      1435 GCGCAGTTCCTCAACATGATCATGATGATGTGTCTCTGCACTGACTTGAATATTCGTC 1494
Qy      355 LeuLeuProMetLysGlyAlaAapProGlnLysSerGlySerAanLysSerLeuAlaLeu 374
Db      1495 GTGGAGGGGTGGTGGACCTTGAGAGATGTGTCTCTGCACTGACTTGAATATTCGTC 1554
Qy      375 ---HisLeuLysLeuAapHisLeuHisLeuProPro-----GlnTyr 387
Db      1555 TCCAAATTTTGGACATCTGTATGTGCTATCTGAGGTAAAGACACCTTGCAGAAAC 1614
Qy      388 LeuHisLeuValAlaAapGlnTyrPheHisAapLysHisSerLeuThrGlnLysAap 407
Db      1615 ATCAAGTTCATGTATACAGACTGG---GCAAGCCGTGACAAACCTGAGACCCCGCGAAA 1671
Qy      408 SerLeuLeuAapLeuSerLysGlyAapValPhePheValValProAlaLeuLysThrAla 427

```

```

Db      1672 ACACTGGTGGCACTCTTCACTGACCAAGTGGGTGAGCCCTCAGTGTGACAGCCGAT 1731
Qy      428 TyrHisArgAapAlaGlyAlaAapProValTyrPheTyrGluLysPheAapHisAapProGln 447
Db      1732 CTGCATGCGCCGCTAGGGCTCGCTCATCTTCTACGCTTCATCAT-----CACTGC 1785
Qy      448 PheGluAapThrLysProAlaPheValLysAlaAapHisAlaAapGluValArgPheVal 467
Db      1786 CAGAGCCTCATGAAAGCTGTGTGTGTGATGATGATGAGTGGGATGAAAGTACCTATGTT 1845
Qy      468 PheGlyGlyAlaPheLeuLysGlyAapHisLeuValMetPheGlu---GlyAlaThrGluLys 486
Db      1846 TTTGGGGTTCCTATGTA---GGCCCCCTGACCTTTCCCTCCAACTTCTCAAGAAAT 1902
Qy      487 GlnLysLeuLeuSerAapGlyLysMetMetLysTyrTrrAlaThrPheAlaArgThrGlyAan 506
Db      1903 GATGTATGTCTCAGTGTGCTGTGTATGATCTTATGACCAACTTTCGCAAGACTGGGGAT 1962
Qy      507 ProAan-----GlyAanAapLeuSer 513
Db      1963 CCAACAAAGCCGGTCCCCCAGACCAAGTTCATTCACCAAGGCCCAACCGCTTGAG 2022
Qy      514 -----LeuTrrProAlaTyrAanLeuThrGluGln---TyrLeuGlnLeuAapLeuAan 530
Db      2023 GAAGTGGCTGTGTCCAATACAAATCCCGAGACCAAGCTTCTTCACTCATCGGGCTGAAA 2082
Qy      531 MetSerLeuGlyGlnArgLeuLysGluProArgValAapPheTrrPheThrLysThrLysPro 550
Db      2083 CCAAGGTTCGAGATCATTTACCGGGCCCATTAAGGTGGCTTTGGAAACATCTGTGGCC 2142
Qy      551 LeuLeuSerAlaSerAapMetLeuHis 560
Db      2143 CACCTATACAAACCTGCATGACATTTTCAC 2172

RESULT 10
US-09-949-016-1017
; Sequence 1017, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OR INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0101307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1017
; LENGTH: 4960
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1017

Alignment Scores:
Pred. No.: 5,996-83
Score: 789.50
Percent Similarity: 51.32%
Best Local Similarity: 33.72%
Query Match: 25.64%
DB: 4
Gaps: 22

US-10-023-515-2 (1-581) x US-09-949-016-1017 (1-4960)
Qy      41 ThrArgLeuGlyTrrPheGlnGly-----LysGlnValThr-----ValLeuGlySer 56
Db      838 ACCAATTTCGAAAGATTAAGAGGATTAAGAAAGAACTCATATTAATTTTGGGG--- 894

```

```

Qy 57 ProVal1AsnVal1PheLeuGlyVal1ProPheAla1a1aProProLeuGlySerLeu 76
Db 895 -----CGTTATTCAATTCTTGGGGTTCCATTGACGCCCAACAAAGGGAAGCT 948
Qy 77 ArgPheThrAsnProGlnProAlaSerProTrrAsnLeuGlyAlaThrSerTyr 96
Db 949 CGTTTACAGCCCTCAAGAACCAACCATCTCCCTGTCAGATATCAAGAAATGCCCAATTT 1008
Qy 97 ProAsnLeuGlyLeuGlnAsnSerGluTrrLeuLeuLeuAspGlnH1metLeuVal 116
Db 1009 GCTCTGTGTGTCCCAAGAT-----ATCATTTGATGGCAGATTGCCAAGATGC 1056
Qy 117 H1strProLysPhe-----GlyVal 123
Db 1057 ATGCTCTCTGTGTGTTACTAATACTTGATGTGGTTTCAATATGTGCAACACAG 1116
Qy 124 SerGluAspGlyLeuTyrLeuAsnH1eTrrAlaProAlaH1s-----AlaAspThr 140
Db 1117 AGCGAAGACTGCGCTATTATAATATATATGTCGCCAGTGAAGATGATATTCGGGACAT 1176
Qy 141 GlySerLysLeuProValLeuVal1TrrPheProGlyGlyAlaPheLysThrGlySerAla 160
Db 1177 GGGGGTCCCAACCAAGATAGGTATATCCATGTGGCTCATATATGGAAGGATCTGGA 1236
Qy 161 SerL1PheAspGlySerAlaLeuAla1a1aTrrGluAspVal1eVal1Val1Gln 180
Db 1237 AATTTATATGATGAAAGTCTCTGGCAAGTTATGGCAATGATGATCTCATCAAGTCAAC 1296
Qy 181 TrrArgLeuGlyLePheGlyPhePheThrTrrTrrAspGlnH1a1aProGlyAsnTrr 200
Db 1297 TATCACTTGGAGTACGCTGTTTCTTGAGTACAGCGCATCAAGGCTCAAAAGGGAAGTAT 1356
Qy 201 AlaPheLysAspGlnVal1Ala1aLeuSerTrrVal1GlnLysAsnH1eGluPheGly 220
Db 1357 GCACTCTCTTGATCTCATCAAGCTTTAAGATGAGTACAGTGAAGAAATGGAATCTTTGGT 1416
Qy 221 GlyAspProSerSerVal1ThrL1PheGlyGlySerAlaGlyAla1eSerVal1SerSer 240
Db 1417 GGTGACCCCTTAAGATCACTGTTTGGATCTGGGTGCTGGGGTTCATGTCTCAAGCTG 1476
Qy 241 LeuLeuLeuSerProMetAla-----LysGlyLeuPhe 251
Db 1477 CTGACTTATCTCCATTATCTGAAGTAAACGTTGAGACAAATTCAACAAAGACTTTT 1536
Qy 252 H1eLysAla1e1metGluSerGlyVal1Ala1e1e1eProTrrLeuGlnAlaH1sAspTrr 271
Db 1537 CAACGAGCAATAGCTCAAAAGTGAACAGCCCTTTCACAGTGG-----GCTGTTAGTTT 1590
Qy 272 GluLysSerGluAspLeuGlnVal1AlaH1sPheCysGlyAsnAsnAlaSerAspSer 291
Db 1591 CAACCTGCAAATATGCTAGATGTTGGCCACAAAGTTGGTGCAGATGTTCAATAC 1650
Qy 292 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer--- 310
Db 1651 GTAGAGTTAGTGAATGCTCAAGAAAGACCTTCAAGAAAGCTTGTGACCAAGATATT 1710
Qy 311 GlnLysThrLys-----SerPheThrArgVal1AlaAspGlyAlaPhePheProAsn 327
Db 1711 CAACGAGCTCGATACACATACCTTGTGACCTGTGATTTGATGATGATGATATACAGAC 1770
Qy 328 GluProLeuAspLeuSerGlnLysAlaPheLysAla1eProSerL1e1eGlyVal 347
Db 1771 GACCCCGAGATTTATAGAGCAAGAGAGATTTTCAACTAGATATATATGTTAGAGATG 1830
Qy 348 AsnAsnH1sGlu---CysGlyPheLeuLeuProMetLysGluAlaProGluLeuLeuSer 366
Db 1831 AACCAAGGGGAAGGTTAAATTTGTTGAAATATAGTAAAGATGAGATGATATATCA 1890
Qy 367 GlySerAsnLysSerLeuAlaLeu---H1sLeuL1eGlnAsnH1eLeuH1s1eProPro 385
Db 1891 GCTAGTGAATTTGACTTGTCTTCAAAATTTTGTGATATATGATGATATCTGAA 1950
Qy 386 -----GlnTrrLeuH1sLeuVal1AlaAsnGluTrrPheH1sAspLys 399

```

```

Db 1951 GCGAAAGATGTTTGAGAGAAACCATTAAGTTCAATGATATCACTGCG---GCTGACCGT 2007
Qy 400 H1sSerLeuThrGlnL1eArgAspSerLeuLeuAspLeuGlyAspVal1PhePheVal 419
Db 2008 CATAACTCTGAACCAAGAAAGACATTAAGTCTGCTTTGTTTACCGACCATCAAGGGGTG 2067
Qy 420 ValProAlaLeuL1eThrAlaArgTrrH1sArgAspAlaGlyAlaProVal1TrrPheTrr 439
Db 2068 GCACCAAGCTTAGCCACAGCGGATCTTCAAACTTGGTTCCCTACCTACGATCTTCAAT 2127
Qy 440 GluPheArgH1sArgProGlnCysPheGlyAspThrLysProAlaPheVal1LysAsp 459
Db 2128 GCCTTTTACAT-----CATTCGCAACAGATCAAGTTCCAGCTTGGGCTGATGACAGC 2181
Qy 460 H1sAlaAspGlyVal1ArgPheVal1PheGlyGlyAlaPheLeuLysGlyAspL1eVal1met 479
Db 2182 CACGAGACGAGGTTCCCTATGATCTGGGAATCCCATGATTT---GGCCCTACAGAGTTA 2238
Qy 480 PheGlu---GlyAlaThrGluGluGluLysLeuLeuSerArgLysLysMetLysTrrTrr 498
Db 2239 TTCTCTGCAATTTCTCCAAAATGATGTGATGCTGAGTGCAGTTGTTATGACATATCTGG 2298
Qy 499 AlaThrPheAlaArgThrGlyAsnProAsn-----GlyAsnAspLeuSerLeu--- 514
Db 2239 ACAATTTTGTCTAAACGTGTCGACCAATCAACACGTCCTCAAGACAGCAAAATTCAT 2358
Qy 515 -----TrrProAlaTrrAsnLeuThrGluGln 523
Db 2359 CATACCAAAACCAACCGTTTGAAGAGTACAGTACAGATATTTCCAGAAAGACCA 2418
Qy 524 ---TrrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArgVal 542
Db 2419 CTATTATCTCAATATGATGATTAACCAAGAGTTAAAGAACTTACAGGCCCAATAGGTG 2478
Qy 543 AspPheTrrThrSerThrL1eProLeuL1eLeuSerAlaSerAspMet-LeuH1sSerPrr 562
Db 2479 AACCTGTGTTGAAGTGTGATCTCATGTGCATATATCAATGACATTTCCAGATAC 2538
Qy 562 GluLeuSer-----SerLeuThrPheLeuSerLeuLeu----- 572
Db 2539 TCTACACACATAAAGTCCATCACTGACATCACTTTCAGACCTTACGAGAAAAATTC 2598
Qy 573 -----GlnProPhe 575
Db 2599 GTACCTGTACGTCAGCTTTC 2620

```

RESULT 11
 US-09-949-016-2516
 Sequence 2516, Application US/09949016
 Patent No. 6812319
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: C1001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 PRIOR FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2516
 LENGTH: 2862
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-2516

Alignment Scores:

Pred. No.: 1.78e-82 Length: 2862
 Score: 782.00 Matches: 199
 Percent Similarity: 46.65% Conservative: 114
 Best Local Similarity: 29.66% Mismatches: 224
 Query Match: 25.40% Indels: 134
 DB: 4 Gaps: 20
 US-10-023-515-2 (1-581) x US-09-949-016-2516 (1-2862)

```

QY      12  TrpCybRhepHeuLeuGlnProLeuLeuGly-----HisArg 25
DB      256  TGGTGGCTTGGACGACCTGCTCTTACATGCTGGGACCTGTAGTCCCTCCGCGGAGAC 315
QY      26  GlnTrpGlySerThrGlyPro-----                32
DB      316  ATGTGGGTGGGGGCTTGGCCCGCCCTGCTCCCTAGACCCCAAGCCAGGTGGCAAG 375
QY      33  -----SerAlaGlu 35
DB      376  AGCCTGTGCTTACCCCTGTGGTTCCTCAGTTTGGCGCTGAGGGCCAGTACCCAGGCCCA 435
QY      36  GlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValSerGlnVal----- 51
DB      436  GCACCCACAGTCAACACTACTTGGGAAAGCTAAAGGGGTGCCAGTACCACTGCCAGT 495
QY      52  ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 71
DB      496  GAGATCTGTGGG-----CCTGTGGACCAATACCTGGGGGTGCTTACCGAGCTGCC 546
QY      72  ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 91
DB      547  CCGATGGGGGAGAAAGCTTTCCTGCCCCCTGAAACCAACCCCATCTGTGGGCAATCCGG 606
QY      92  GluAlaThrSerTyrrProAsnLeuCyLeuGlnAsnSerGlu----- 105
DB      607  AAGCGCACACACTTTCCTCCAGGTGGCCCCCAGAAACATCCACACAGCTGTGCCCAAGTC 666
QY      106  -----TrpLeuLeuLeuAsnGlnHisMetLeuValHisTyrrProValPhe 121
DB      667  ATGCTGCCGCTGTGTTCACTGCCCACTTGATATCGTACCTTACATCCAGAG--- 723
QY      122  GlnValSerGlnAspCyLeuTyrrLeuAsnIleTyrrAlaProAlaHis----- 137
DB      724  ---CCCAAGAGAGACTGTCTCTACCTGAAAGCTTATGTGGCCAGCGAAGATGTAAAGCG 780
QY      137  ----- 137
DB      781  ATTTCCAAAGATGCGCCCGAAAGCCCAACAAGAAATTTGTAGGAAGAGATCGCGC 840
QY      138  -----Ala 138
DB      841  GCTAAGAAACAGGGGAGAGACTTACCGGATATGACGGGGATGAAGATGAAGATCCGG 900
QY      139  AspThrGlySerTyrrLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGly 158
DB      901  GACAGTGTGGCTTAAA---CCCGTATGTGTCAATTCACCGAGAGCTCTTACATGAAAGGG 957
QY      159  SerAlaSerIlePheAspGlySerAlaLeuAlaAlaTyrrGluAspValLeuValVal 178
DB      958  ACAGGCAACATGATGATGATGAGGACATCTCCGCAAGTTATGGCAATGTCAATCATCACC 1017
QY      179  ValGlnTyrrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHisAlaProGly 198
DB      1018  CTCAACTATGCGGGTGGAGTGTAGGTTTCTCTGAGTACTGGAAGTCAAGGCTCCAGAGGC 1077
QY      199  AsnTrpAlaPheLeuValAspGlnValAlaAlaLeuSerTrpValGlnValAsnIleGluPhe 218
DB      1078  AACATATGGGCTCTTGAACAGATCCAGGCTCCGCGGGGTGAGCGAGATATATGCTTCC 1137
QY      219  PheGlyValAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerVal 238
DB      1138  TTGGGGGAGAACCCCGCGGATCATGTCTTGGCTGGGGCATGTGTGATCTTGTGCTC 1197
  
```

```

QY      239  SerSerLeuIleLeuSerProMetAlaValGlyLeuPheHisIleValAlaIleMetGluSer 258
DB      1198  AGCCTCTTCACTGATCATCATCATCAGAGGAGACTTTCCAGAGAGCCATCATCCAAAGT 1257
QY      259  GlnValAlaIleIleProTyrrLeuGlnAlaHisAspTyrrGlyValSerGluAspLeuGln 278
DB      1258  GGCCTGCTGTGTCCAGGCTGG---GCTGTGAACCTTCAACCAAGTGAATGACCAAGCAGC 1311
QY      279  ValValAlaHisPheCyGlyValAsnAsnAlaSerAspSerGluAlaLeuLeuArgCyLeu 298
DB      1312  CTGTGGCAGAACAAAGGTGGCTGTAAATGTGCTGACACCGTGAATGTGGAGACTGTCTT 1371
QY      299  ArgThrIleProSerIleValLeu-----LeuThrLeuSerGlnTyrrLeuVal 314
DB      1372  CCGCAAAAGAGTGGCCAGAGAGCTGTGAGAGCAGGACATCCAGCCGCGGTACAGGTG 1431
QY      315  SerPheThrArgValValAspGlyValaPhePheProAsnGluProLeuAspLeuSer 334
DB      1432  GCCTTTGGCCCTGTGATGTGATGTGATGTCAATTCTGTAGTACCTTGAGATCTTCATGAG 1491
QY      335  GlnValaPheLeuValaIleProSerIleIleGlyValaAsnHisGlu---CyGly 353
DB      1492  CAGGGCCAGTTCCTCACTATGATCATCATGTAGTGTCAACCAAGGCGAGGGGTCTCAAG 1551
QY      354  PheLeuLeuProMetIleValaIleProGlnIleLeuSerGlySerAsnIleVala 373
DB      1552  TTTGTGAAAGGGGTGGAGAGCTTGAAGATGTGTCTGTGGCACTGACTTGAATATTCC 1611
QY      374  Leu---HisValIleGlnAsnIleLeuHisIleProPro-----Gln 386
DB      1612  GTCTCCAAATTTTGGACAAATCTGTATGCTATGCTTGAAGGTAAAGACACCTGCCAGAG 1671
QY      387  TyrrLeuHisLeuValAlaAsnGluTyrrPheHisAspIleHisSerLeuThrGluLeuArg 406
DB      1672  ACCATCAAGTTCATGTATACAGACTGG---GCAGACCCGTGACAACTGAGACCCCGCT 1728
QY      407  AspSerLeuLeuAspLeuLeuGlnGlyAspValPhePheValValProAlaLeuIleTrpAla 426
DB      1729  AAAACACTGGTGACCTTTCATCACTGACACCAAGTGGGTGAGGCCCTCAAGTGTGACAGCC 1788
QY      427  ArgTyrrHisArgAspAlaGlyAlaProValTyrrPheTyrrGluPheArgHisArgProGln 446
DB      1789  GATCTGATGTCGCCGCTACGAGCTGCTGCTTCACTTCTACGCGCTTATCAT-----CAC 1842
QY      447  CyPheGluAspThrIleAspProAlaPheValIleValAspHisAlaAspGluValaArgPhe 466
DB      1843  TGCCAGAGCTCATGAAAGCTGTGTGTGTCAGATGACGTCATGAGGGAGTGAAGTACCCAT 1902
QY      467  ValPheGlyGlyAlaPheLeuValAspIleValMetPheGlu---GlyAlaThrGlu 485
DB      1903  GTTTTGGGGTTCCTATGGTA---GGCCCACTGACCTTTTCCCTGCAACTTCTCCAAAG 1959
QY      486  GlnGluValLeuLeuSerTyrrGlyHisMetMetIleTyrrTrpAlaThrPheAlaArgThrGly 505
DB      1960  AATGATGTTATGTCTGAGTGTGTGTCATGACCTTATGTGACCAACCTTGGCCAAAGCTGGG 2019
QY      506  AsnProAsn-----GlyAsnAspLeu 512
DB      2020  GATCCCAACAAACCGGCTGCCCGCAGAGACCAAGTTCACTTACACCAAGGCCACCGCTTT 2079
QY      513  Ser-----LeuTrpProAlaTyrrAsnLeuThrGluGln---TyrrLeuGlnLeuAspLeu 529
DB      2080  GAGGAAGTGGCTGTGTCCAAATATCAATCCCGAGACAGCTTCACTTACATCGGGCTG 2139
QY      530  AsnMetSerLeuGlyGlnArgLeuValGluProArgValaAspPheTrpThrSerThrIle 549
DB      2140  AAACCAAGGCTCCGAGATCATTTACGGGCACTAAAGTGGCTTTTGGAAACATCTGTGTG 2199
QY      550  ProLeuIleLeuSerAlaSerAspMetLeuHis 560
DB      2200  CCCCACTATACAACTCATGACATGTTCAC 2232
  
```

RESULT 12

US-09-491-356C-5
Sequence 5, Application US/09491356C
Patent No. 656061
GENERAL INFORMATION:
APPLICANT: Philibert, Robert A.
APPLICANT: Glines, Edward I.
APPLICANT: Delist, Lynn
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCT4 REGION OF XQ13
FILE REFERENCE: 9465.6US11
CURRENT APPLICATION NUMBER: US/09/491,356C
CURRENT FILING DATE: 2000-01-26
PRIOR APPLICATION NUMBER: PCT/US99/09365
PRIOR FILING DATE: 1999-04-29
PRIOR APPLICATION NUMBER: 60/083,465
PRIOR FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 4233
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4171)..(4171)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4179)..(4179)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4181)..(4181)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4197)..(4197)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4228)..(4228)
OTHER INFORMATION: n is not determined
US-09-491-356C-5
Alignment Scores:
Pred. No.: 2,476-81 Length: 4233
Score: 775.00 Matches: 188
Percent Similarity: 50.42% Conservative: 113
Best Local Similarity: 31.49% Mismatches: 220
Query Match: 25.17% Indels: 76
DB: 4 Gaps: 17
US-10-023-515-2 (1-581) x US-09-491-356C-5 (1-4233)
QY 29 LysThrGlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGly 48
DB 506 AGCGGCAAGTACCCAGGAGCCCGGAGCCAGCCAGTCACTGTTGGAGAGTAAAGGGT 565
QY 49 LysGlnVal-----ThrValLeuGlySerProValProValAsnValPheLeu 64
DB 566 GCCCAGATGACCACTGCCAGATGATCTGGGG-----CTGTGACCAATACTGTG 616
QY 65 GlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsnProGlnProAla 84
DB 617 GGGGGGCCCCATGAGAGCTCCCGGATCGGCGAGAAAGTTTCTGCCCCCTGAACCAACC 676
QY 85 SerProTrpAspAsnLeuArgGluAlaThrSerTrpProAsnLeuCysLeuGlnAsnSer 104
DB 677 CCATCTCTGTGGGAGATCGGAGACGCAACAACATTTCCCGGAGTGGCCCCCAAGAACTC 736
QY 105 Glu-----TrpLeuLeuLeuAspGlnIleMetLeu 114
DB 737 CACACAGCTGTGCGGAGATCACTGTCGGGTCTGTTCACTGCACATTTGGATATCGTC 796
QY 115 LysValHisTrpProLysPheGlyValSerGluAspCysLeuTrpLeuAsnIleTrpAla 134

DB 797 GCTACTTACATCCAGAG-----CCCAAGAAAGTCTCTTCACTGAACGTATATGTG 850
QY 135 ProAlaHisAlaPheThrGlySerLys----- 143
DB 851 CCGACGAGAGTGGATTCGGCGCTTAAGAAACAGGCGAGACTTAGCGGATATGACGGG 910
QY 144 -----LeuProValIleValTrpPheProGly 152
DB 911 GATGAAGATGAAGACATCCGGAGACGTGTGCTAAACCGTCATGTGTCTACATCCACGA 970
QY 153 GlyAlaPheLysThrGlySerLysSerIlePheAspGlySerAlaLeuAlaTrpGlu 172
DB 971 GGCTCTTACATGAAGGAGACAGGCAACATGATGAGCAGATCTTGCGCCAGTATGAGC 1030
QY 173 AspValIleValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrTrp 192
DB 1031 AATGTCATAGTCATCACTCACTCACTGAGGTGGAGTGAAGTATGATTTCTGATCTGGA 1090
QY 193 AspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaIleSerTrpVal 212
DB 1091 GATTCAGCTGCCAAGGAGCACTATGAGCTCTTACCAAGATCCAGGCCCTCGCGGTG 1150
QY 213 GlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGluSer 232
DB 1151 AGCGAGATATTCCTCTTCCGGGGAGACCCCGCGAGTACATGCTTTGGCTCGGAGC 1210
QY 233 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPheHis 252
DB 1211 ATTTGTGATCTGTGGTCACTCTCACTGTTGTCATCTCACTGAGAGGACTTTTCCAG 1270
QY 253 LysAlaIleMetGluSerGlyValAlaIleIleProTrpLeuGluAlaHisAspTrpGlu 272
DB 1271 AGAGCATCATCAAAAGTGGCTGTGCTGTCCAGCTGG-----GCTGTAACTCAAA 1324
QY 273 LysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSerGlu 292
DB 1325 CCAAGTGAAGTACACAGGCTGTGSCAGACAAAGTGGCTGAATGCTGTGACACCGTG 1384
QY 293 AlaLeuLeuArgCysLeuArgThrLysPheSerLysGluLeu-----LeuThr 308
DB 1385 GATATGTGTGACTGTCTTCCGCAAAAGTGCCAAGGAGCTGTGACAGCAGACCACTCAG 1444
QY 309 LeuSerGlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGlu 328
DB 1445 CCAAGCCGCTTACCAAGTGGCTTTGGCCCTGTGATGATGATGATGATGATGATGATG 1504
QY 329 ProLeuAspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsn 348
DB 1505 CCTGAGATCTCATAGAGACAGGCGAGTCTCTCAACTGATGACATGATGATGATGATGATG 1564
QY 349 AsnHisGlu-----CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGly 367
DB 1565 CAGGCGAGGAGTCTCAAGTTTGTGAGAGGGTGTGGAGCCCTGAGAGATGTTCTCTGGC 1624
QY 368 SerAsnLysSerLeuAlaLeu-----HisLeuIleGlnAsnIleLeuHisIleProPro 385
DB 1625 ACTGACTTGAATATCTCGCTCTCAATTTTGTGCAATCTGTATGAGTATGAGTATGAGGT 1684
QY 386 -----GlnTrpLeuHisLeuValAlaAsnGlnTrpPheHisAspLysHis 400
DB 1685 AAGACACCCCTCGAGAGACCATCAAGTTCATGTATACAGCTG---CGAGACCGGTGAC 1741
QY 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyLysValPhePheValVal 420
DB 1742 AACCTGAGACCCCGCGTAAACACTGGTGGCAGCTTCACTGACACCAAGAGGGTGGAG 1801
QY 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGlu 440
DB 1802 CCTCAGTGTGACAGCGGATGTGATGCTCGGCTTACGAGCTGCTGCTCACTTCTTACGCC 1861
QY 441 PheArgHisArgProGlnCysPheGluAspThrLysPheAlaPheValLysAlaAspHis 460

```

Db      1862 TTCTATCAT-----CACTGCCAGAGCCTCATGAAGCCTGTGTCATGACGCTCAT 1915
Qy      461 AlaAspGluValArgPheValPheGlyGlyAlaIshLeuIysGlyAspIleValMetPhe 480
Db      1916 GGGGATGAAGTACCTTATGTTTGGGGTTCCTATGTA---GGCCCCACTGACCTTTTC 1972
Qy      481 Glu---GlyAlaThrGluGluGluValLeuLeuSerArgIshMetLeuIshTyrTPAla 499
Db      1973 CCTGCAACTCTTCACGAATGATGCTCATGCTCATGCTCATGACCTTATGACC 2032
Qy      500 ThrPheAlaArgThrIshIshAsnProAsn----- 508
Db      2033 AACTTCCCAAGACTGGGGAATCCCAACAGCCGCTCCCGAGACACCAAGTTCAATTCAC 2092
Qy      509 -----GlyAsnAspLeuSer-----LeuTyrProAlaTyrAsnLeuThrGluGln--- 523
Db      2093 ACCAAGGCCAACCGCTTTGAGGAAGTGGCCTGCTCCAAATACATCCCGAGACAGCTTC 2152
Qy      524 TyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuIshGluProArgValAsp 543
Db      2153 TACCTTCACATCGGCTGGAACCAAGGGTCCGAGATCATATCCGGGCGCACTAAGTGGCC 2212
Qy      544 PheTyrThrSerThrIshProLeuIshLeuSerAlaSerAspMetLeuIsh 560
Db      2213 TTTTGAACATCTGTGTGCTCCCACTTATACACCTGCAATGACATGTTCCAC 2263

```

RESULT 13
US-07-732-962A-1
Sequence 1, Application US/07732962A
Patent No. 5248604
GENERAL INFORMATION:

```

APPLICANT: Fischer, Melr
TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
NUMBER OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
CORRESPONDENCE ADDRESSES: 2
ADDRESSES: John P. White, Esq.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07732,962A
FILING DATE: 19910722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
US-07-732-962A-1

```

Alignment Scores: 1.02e-80 Length: 1845
Pred. No.:

Score: 764.50
Percent Similarity: 49.83%
Best Local Similarity: 33.72%
Query Match: 24.83%
DB: 1
Gaps: 22

US-10-023-515-2 (1-581) x US-07-732-962A-1 (1-1845)

```

Qy      1 MetProGlnGlyLeuThrSerSerAlaSerGlnTyrCysPhePheLeuIleLeuGlnPro 20
Db      41 CTTCCTCCACTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 79
Qy      21 LeuLeuGlnIshArgGlnTyrGlyIshThrGlyProSerAlaGluGlyProGlnArgAsn 40
Db      80 -----GAGAGTGGGG----- 91
Qy      41 ThrArgLeuGlyTyrPheGlnGlyLeuGlnValThrValLeuGlySerProVal----- 58
Db      92 CTGAGGGCGCGGAGAGATGACAGCTG-CTGCTGACGGTGGTGGGGCGCGGCGGCGGCG 150
Qy      59 -----ProValAsnValPheLeuGlyValProPheAlaAla 70
Db      151 ATTGCGCTGAAGACCCCGGGGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 210
Qy      71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAsnProVal 90
Db      211 CCACCCATGGGACCCCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 270
Qy      91 ArgGlnAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGlnTyrLeuLeuAsp 110
Db      271 GTAGACGCTACCAACCTTCCAGAGTGTCTTACCAATATGTGACACCTTATCCAGAGT 330
Qy      111 GlnIshMetLeuIshValIshTyrProIshPheGlyValSerGlnAspCysLeuTyrLeu 130
Db      331 TTGAGGGCACCCAGATGTGGAACCCCAACCGTGAAGTGGAGAGTGGAGTGGAGTGGAGT 390
Qy      131 AsnIleTyrAlaProAlaIshAlaAspThrGlySerTyrLeuProValLeuValTyrPhe 150
Db      391 AAGGTGTGACACATATACCCCGGCTTACA--TCCCCACCCCTGTCTCTCTCTCTCTCTCT 447
Qy      151 ProGlyGlyAlaPheIshThrGlySerAlaSer-----IlePheAspGlySerAlaLeu 168
Db      448 TATGGGGGTGGCTTTCACAGTGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 507
Qy      169 AlaAlaTyrGluAspValLeuValValAlaGlnTyrArgLeuGlyIshPheGlyPhe 188
Db      508 GTACAGCGCGAGAGACTGTGCTGTGTCTCATGACTACCGGGGTGGAGCCCTTGGCTTC 567
Qy      189 PheThr--ThrTyrAspGlnIshAlaProGlyAsnTyrAlaPheIshAspGlnValAla 207
Db      568 CTGGCCCTGCGGGGAGCCGAGAGGCCCGGCAATGTGGTCTCTCTCTCTCTCTCTCTCTCT 627
Qy      208 AlaLeuSerTyrValGlnIshAsnIleGlnPhePheGlyIshAspProSerSerValThr 227
Db      628 GCCCTGACGTGGGTGAGGAGAACGTGACAGCTTCTGGGGGTGACCCGACATCATGTAAG 687
Qy      228 IlePheGlyIshSerAlaGlyAlaIshSerValSerSerLeuIleLeuSerProMetAla 247
Db      688 CTGTTGGGGAAGCGCGGAGCCCGCTCTGCTGGGACATGACCTGTCTCTCTCTCTCTCTCT 747
Qy      248 IysGlyLeuPheIshValAlaIshMetGlnSerGlyValAlaIleIleProTyrLeuGlu 267
Db      748 CGGGGCTGTTCACAGAGCGCTGTGACAGCGGTGCCCAATGACCTCTGGGCCACG 807
Qy      268 AlaIshAspTyrGlnIshSerGlnAspLeuGlnValAlaIshIshPhe-----Cys--- 284
Db      808 GTGGGCAATGGAGAGGCCCGCTGCGAGGGCCAGCAGAGCTGGGCCACTTGTGGGCTGTCT 867
Qy      285 -----GlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThrIshPro 302
Db      868 CCAAGCGGCACTGTGTGGGAATGACACAGAG--CTGTAGCTCTCTCTCTCTCTCTCTCTCT 924
Qy      303 SerIshGlnLeu-----ThrLeuSerGlnIshThrIsh----- 314

```



```

Qy 91 ArgGluAlaThrSerGlyProValLeuCYsLeuGlnAsnSerGluTrpLeuLeuAsp 110
Db 365 GTAAGCGTACAGCTTCCAGAGTGTCTGTACCAATATGTGGACACCTTATCCAGAGT 424
Qy 111 GlnH1MetLeuLYsValH1sTrpOlyvPheGlyValSerGluAspCYsLeuTrpLeu 130
Db 425 TTTGAGGCGACCGAGATGTGGAACTCCAGCCGTGAGAGAGAGACTGCTGTACCTC 484
Qy 131 AsnLeuTrpAlaProAlaH1sAlaAspH1sLeuSerGlyvLeuProValLeuValTrpHe 150
Db 485 AACGCTGAGACCACTACCCCGGCTTACA---TCCCCACCCCTGTCTGTCTGGATC 541
Qy 151 ProGlyGlyAlaPheLYsThrGlySerAlaSer-----IlePheAspGlySerAlaLeu 168
Db 542 TATGGGCTGGCTTCTACAGTGGGCTCTCTCTTGGAGCGTGAAGATGGCCGCTTCTG 601
Qy 169 AlaAlaTrpGluAspValLeuValValValGlnTrpArgLeuGlyTrlePheGlyPhe 188
Db 602 GTACAGCGCCGAGAGACTGTCTGTCTGCATGAATACCGGCTGGAGCCCTTGGCTTC 661
Qy 189 PheThr---ThrTrpAspGlnH1sAlaProGlyAsnTrpAlaPheLYsAspGlnValAla 207
Db 662 CTGCGCCCTGCGCGGAGCCGAGAGCCCGGCGCAATGTGGCTTCTGTGAATCAGAGCTG 721
Qy 208 AlaLeuSerTrpValGlnLYsAsnIleGluPhePheGlyGlyAspProSerSerValThr 227
Db 722 GCCCTGCACTGGGTCCAGAGAACCTGGCAGCCTTCCGGGGGTGACCCCACTCACTGACG 781
Qy 228 IlePheGlyGluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAla 247
Db 782 CTGTTTGGGAGAGCGCGGAGCCCGCTGGTGGCGATGCACCTGTCCCGCCGACG 841
Qy 248 LyseGlyLeuPheH1sLYsAlaIleMetGluSerGlyValAlaIleIleProTrpLeuIle 267
Db 842 CGGCGCTGTTCACAGCGCGCTGTCTGCAGAGCGGTGCCCAATGAGCCCTGGCCAGC 901
Qy 268 AlaH1sAspTrpGlyLYsSerGlyvAspLeuGlnValAlaH1sPhe-----Cys--- 284
Db 902 GTGGGCAATGGAGAGAGCCCGTCCGAGGGCCAGCGAGCTGGCCACTTGTGGGCTGTCT 961
Qy 285 -----GlyAsnAsnAlaSerAspSerGlyvAlaLeuLeuAspArgCYsLeuAspTrp 302
Db 962 CCAGCGCGCACTGGTGGAGATGACACAGAG---CTGTAGGCTGTCTGACACAGACCA 1018
Qy 303 SerLYsGluLeuLeu-----ThrLeuSerGlyvLeuThrLYs----- 314
Db 1019 GCGCAGCTCTGTGTGAACCCAGAAATGGCACTGTCTCTCAAGAAAGCTTCCGAGTTC 1078
Qy 315 SerPheThrArgValAlaAspGlyValaPhePheProAsnGluProLeuAspLeuSer 334
Db 1079 TCTTTCGTGCTGTGTGTAGTGAAGATTTCTCTCACTGACACCCCGAGAGCCCTCATCAAC 1138
Qy 335 GlnLYsAlaPheLYsAlaIleProSerIleIleGlyValaAsnAsnIleGluCYsGlyPhe 354
Db 1139 GCGGAGAGACTTCCAGCGCTGAGGTGCTGGTGGTGTGTGAAGATGAGAGGCTCGTAT 1198
Qy 355 LeuLeuProMetLYsGlyAlaProGluIleLeuSerGlySerAspLYsSerLeu----- 372
Db 1199 TTTCTG---GTTTACGGGCGCCCAAGC---TTCAAGAAAGACAGAGTCTTCTCATCAGC 1252
Qy 373 AlaLeuH1sLeuIleGlnAlaAsnIleLeuH1sIleProProGlnTrpLeuH1sLeuValAla 392
Db 1253 CCGGCGCAAGTTCCTGGCGGGGTGCGGGTTCCTCCCAAGTAAAGTGAACCTGGCAGCC 1312
Qy 393 -----AsnGlyTrpPheH1sAspLYsAspLYsAspLYsSerLeuThrGluIle 405
Db 1313 GAGGCTGTGTCTGTATTAACAGACTGTGTGATCCAGAGACCG---GCAGCGCTG 1369
Qy 406 ArgAspSerLeuLeuAspLeuGluLYsAspValPhePheValValProAlaLeuIleThr 425
Db 1370 AGGAGCGCTGAGAGATGTGTGGCGAGCAACAATGTCTGTGTGCCCTGGCCAGCTG 1429
Qy 426 AlaArgTrpH1sArgAspAlaGlyValaProValTrpPheTrpGluPheArgH1sArgPro 445

```

```

Db 1430 GCTGGGCGACTGTGCTGCCAGAGGTGCCCGGAGTCTACGCTGTCTTGAACACCGTCT 1489
Qy 446 GlnCYsPheGluAspTrpTrpLYsProAlaPheValLYsAlaAspH1sAlaAspGluValArg 465
Db 1490 TCCAGCGCTC-----TCTGTGCCCTGTGTGATGGAGGCTGCCCAAGGCTTACAGATCGAG 1543
Qy 466 PheValPheGlyGlyAlaPheLeuLYsGlyAspIleValMetPheGluGlyAla----- 483
Db 1544 TTCACTTTGGG-----ATCCCGCTGGAGCCCTTCGAAAC 1579
Qy 484 ---ThrGluGlnGlyLYsLeuLeuSerArgLYsMetMetLYsTrpAlaIleThrPheAla 502
Db 1580 TACACGGCAGAGAGAAATCTTCCGCCAGAGCTGATGCATATCTGGGCCCACTTTC 1639
Qy 503 ArgTrpGlyAsnProAsn-----GlyAsnAspLeuSerLeuTrpProAlaTrpAsn 519
Db 1640 CGCACAGGGAGTCCCAATGAGCCCGGAGACCCCAAGGCCCAACATGGCCCGCTACAGC 1699
Qy 520 Leu---ThrGluGlnTrpLeuGlnLeuAspLeuAsn---MetSerLeuGlyGlnArgLeu 537
Db 1700 GCGGGGCTCAGCAGTACGTTAGTGTGAGCTTGGCGCGCTGGAGAGTCCGGCGGGGCTG 1759
Qy 538 LyseGluProArgValaAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAsp 557
Db 1760 CGCGCGCAGGCTGCGCTTGTGAACCGCTTCTCTCCCAATATGTCTCAGCGCCACCGAC 1819
Qy 558 MetLeu 559
Db 1820 ACGCTC 1825

```

Search completed: June 15, 2005, 13:34:22
Job time : 331 secs

This Page Blank (uspto)

Qy	1	CCAGCGCTCCGAAAAACAAGCCTTGAGAGCAATGTGGAGTAAGCAATGTAAATAAACGA	60
Db	1	CCAGCGCTCCGAAAAACAAGCCTTGAGAGCAATGTGGAGTAAGCAATGTAAATAAACGA	60
Qy	61	TTTAAAAATTATCTTAATAAAGTACGAATCCCAATGCCACAGGAACTTAATCTTCACTG	120
Db	61	TTTAAAAATTATCTTAATAAAGTACGAATCCCAATGCCACAGGAACTTAATCTTCACTG	120
Qy	121	CTTCACAAATGTGCTTTTCTCTGATTTCTCAAGCCCTGTGTGGACAACAACATGTGGGAA	180
Db	121	CTTCACAAATGTGCTTTTCTCTGATTTCTCAAGCCCTGTGTGGACAACAACATGTGGGAA	180
Qy	181	AAACTGGGCTTCTGCTGAAGGGGCAAGAGGAAACAACAGCTGGATGTATTAAGGCA	240
Db	181	AAACTGGGCTTCTGCTGAAGGGGCAAGAGGAAACAACAGCTGGATGTATTAAGGCA	240
Qy	241	AGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCCCTTTG	300
Db	241	AGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCCCTTTG	300
Qy	301	CTGCTCCCCCGCTGGGATCCCTGTGATTTACGAACCCGACGCTGCATGCGCCCTGGGATA	360
Db	301	CTGCTCCCCCGCTGGGATCCCTGTGATTTACGAACCCGACGCTGCATGCGCCCTGGGATA	360
Qy	361	ACTTGGAGAGGCACCTCCTACCCCTAATTGTGCGCTCCAGAACTGAGATGCTGCTC	420
Db	361	ACTTGGAGAGGCACCTCCTACCCCTAATTGTGCGCTCCAGAACTGAGATGCTGCTC	420
Qy	421	TGATTAACAACATGCTCCAGGTGATTAACCGAAATTGGAGTGTCAAGAACTGCTCT	480
Db	421	TGATTAACAACATGCTCCAGGTGATTAACCGAAATTGGAGTGTCAAGAACTGCTCT	480
Qy	481	ACCTGAACATCTATGCGCTGCCACGCGCATACAGGCTTCAGGCTCCCGCTTGGTGT	540
Db	481	ACCTGAACATCTATGCGCTGCCACGCGCATACAGGCTTCAGGCTCCCGCTTGGTGT	540
Qy	541	GCTTCCCAAGAGGTGCTTCAAGACTGAGCTGACCTCACTTTTGAATGGGTCCGCTG	600
Db	541	GCTTCCCAAGAGGTGCTTCAAGACTGAGCTGACCTCACTTTTGAATGGGTCCGCTG	600
Qy	601	CTGCGTATGAGACGCTGCTGTGGTGTGTCACAGTACCGGCTAAGAAATTTTGGTTTCT	660
Db	601	CTGCGTATGAGACGCTGCTGTGGTGTGTCACAGTACCGGCTAAGAAATTTTGGTTTCT	660
Qy	661	TCACCAATGGGATCAGATGCTCGGGGAACTGGGCTTCAAGACACAGTGTGCTCTC	720
Db	661	TCACCAATGGGATCAGATGCTCGGGGAACTGGGCTTCAAGACACAGTGTGCTCTC	720
Qy	721	TGCTCTGGGTCCAGAAAGAACATCGAGTTCTTTGGGTGGGACCCCAAGCTTGTGACCATCT	780
Db	721	TGCTCTGGGTCCAGAAAGAACATCGAGTTCTTTGGGTGGGACCCCAAGCTTGTGACCATCT	780
Qy	781	TTGGGAGTCCGCGGAGCCATAAGTGTTCAGTCTTAACTGTCTCCCATGGCCAAAG	840
Db	781	TTGGGAGTCCGCGGAGCCATAAGTGTTCAGTCTTAACTGTCTCCCATGGCCAAAG	840
Qy	841	GCTTATTTCCAAAGCCCATCATGAGAGTGGGTGGCCATCATCTCTTAACCTGAGAGCC	900
Db	841	GCTTATTTCCAAAGCCCATCATGAGAGTGGGTGGCCATCATCTCTTAACCTGAGAGCC	900
Qy	901	ATGATTTATGAAGAAGTAGAGACCTGAGAGTGTGTCATTTCTGTGTGTAAACAATGTGT	960
Db	901	ATGATTTATGAAGAAGTAGAGACCTGAGAGTGTGTCATTTCTGTGTGTAAACAATGTGT	960
Qy	961	CAGACTCTGAGGCCCTGTGAGTGTCTGAGAGCAAAAACCTCCAAAGAGCTGTGACCC	1020
Db	961	CAGACTCTGAGGCCCTGTGAGTGTCTGAGAGCAAAAACCTCCAAAGAGCTGTGACCC	1020
Qy	1021	TCAGCCGAAAAACAAGTCTTCACTGAGTGTGTATGTGTGCTTTCTTCTAATAGC	1080
Db	1021	TCAGCCGAAAAACAAGTCTTCACTGAGTGTGTATGTGTGCTTTCTTCTAATAGC	1080

Qy	1081	CTCTGATCTTATTTGCTCAGAAAGCATTTAAAGCAATTCCTTTCATCATCGAGTCAATA	1140
Db	1081	CTCTGATATCTATTTGCTCAGAAAGCATTTAAAGCAATTCCTTTCATCATCGAGTCAATA	1140
Qy	1141	ACCAAGAGTGGCTTCCGCTGCCCTATGAAAGAGGCTCCCGAATCCTAGTGGCTCCA	1200
Db	1141	ACCAAGAGTGGCTTCCGCTGCCCTATGAAAGAGGCTCCCGAATCCTAGTGGCTCCA	1200
Qy	1201	ACAACTCCCTTGGCCTCCATCTGATGATACAAAACATCTGCACATCCCGCTCAGTATTTGC	1260
Db	1201	ACAACTCCCTTGGCCTCCATCTGATGATACAAAACATCTGCACATCCCGCTCAGTATTTGC	1260
Qy	1261	ACCTGTGGCTAATGAATPACTTCCATGACAAAGCATCCCTGATCTGAAATCCGAGACATC	1320
Db	1261	ACCTGTGGCTAATGAATPACTTCCATGACAAAGCATCCCTGATCTGAAATCCGAGACATC	1320
Qy	1321	TTCTGGAATTGCTGAGATGTGTTCTTGTGGTCCCTGACATGATCAAGCTCGATATC	1380
Db	1321	TTCTGGAATTGCTGAGATGTGTTCTTGTGGTCCCTGACATGATCAAGCTCGATATC	1380
Qy	1381	ACAGAGATGCTGGTGGCACCTGTCTACTTCTATGATTTGGGCAACCGAGCTCAGTGCTTTG	1440
Db	1381	ACAGAGATGCTGGTGGCACCTGTCTACTTCTATGATTTGGGCAACCGAGCTCAGTGCTTTG	1440
Qy	1441	AAGACAGGAAGCCGGCTTTTGTCAAAAGCCGACACGCTGATGAAAGTCCGCTTTGTGGTGG	1500
Db	1441	AAGACAGGAAGCCGGCTTTTGTCAAAAGCCGACACGCTGATGAAAGTCCGCTTTGTGGTGG	1500
Qy	1501	GTGTGCTCTTCTGAAAGGGGACATTTGTAATGTTTGAAGAGCCACGAGAGAGAGAAAT	1560
Db	1501	GTGTGCTCTTCTGAAAGGGGACATTTGTAATGTTTGAAGAGCCACGAGAGAGAGAAAT	1560
Qy	1561	TACTGACCCGGAAGATGATGAAATPACTGGGCTATCCTTCTCGAATCCGGGAAATCTTAATG	1620
Db	1561	TACTGACCCGGAAGATGATGAAATPACTGGGCTATCCTTCTCGAATCCGGGAAATCTTAATG	1620
Qy	1621	GGAACGACTTGTCTGTGGCCAGCTTAATCTGACTGAGCATCTCCAGCTGAGCT	1680
Db	1621	GGAACGACTTGTCTGTGGCCAGCTTAATCTGACTGAGCATCTCCAGCTGAGCT	1680
Qy	1681	TGAACATGAGCCCTCGACAGAGACTAAGAAACCGGGGTGATTTTGGACACAGACCA	1740
Db	1681	TGAACATGAGCCCTCGACAGAGACTAAGAAACCGGGGTGATTTTGGACACAGACCA	1740
Qy	1741	TCCCCCTGATCTGTGTGCTCCGACATGCTCCAGTCTCTTCTTCTTAATCTTCC	1800
Db	1741	TCCCCCTGATCTGTGTGCTCCGACATGCTCCAGTCTCTTCTTCTTAATCTTCC	1800
Qy	1801	TCTCTCTCCCTCAAGCCTTCTTTTCTTTTGTGTCTCTTGAAGATTAATCTTCTGTGAT	1860
Db	1801	TCTCTCTCCCTCAAGCCTTCTTTTCTTTTGTGTCTCTTGAAGATTAATCTTCTGTGAT	1860
Qy	1861	TTTGTGTTCCCTCTCTCCCATTAATTTCTCCGGAATCATTAAGTCTTTCTGAGCTCA	1920
Db	1861	TTTGTGTTCCCTCTCTCCCATTAATTTCTCCGGAATCATTAAGTCTTTCTGAGCTCA	1920
Qy	1921	GCCTCTTCTAATGGGAGTCTTGGCAAAAACAAGCTGCTTGGCTGATATTTATGACTTA	1980
Db	1921	GCCTCTTCTAATGGGAGTCTTGGCAAAAACAAGCTGCTTGGCTGATATTTATGACTTA	1980
Qy	1981	GGAAATGATCCTTAACGAATTCCTTTCACATCAAAAGATGCAATTTGCTTGGAAAGCCA	2040
Db	1981	GGAAATGATCCTTAACGAATTCCTTTCACATCAAAAGATGCAATTTGCTTGGAAAGCCA	2040
Qy	2041	CAAGATTTCTTCAATPAAATTTGGAAAGGGCTGGCTATTAAGTGTCAATAATATGCTTT	2100
Db	2041	CAAGATTTCTTCAATPAAATTTGGAAAGGGCTGGCTATTAAGTGTCAATAATATGCTTT	2100
Qy	2101	TGTAACTCATATGAATAATTAATTCGAATGTAAATATAGAAAAAATTTTTTTTTTTTTTTT	2158
Db	2101	TGTAACTCATATGAATAATTAATTCGAATGTAAATATAGAAAAAATTTTTTTTTTTTTTTT	2158

RESULT 2
US-10-674-636-1
; Sequence 1, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Imaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674,636
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-674-636-1

Query Match 100.0%; Score 2158; DB 18; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGCGTCCGAAAAACAGGCTGAGAGCAATGTGAGTAAGCAATGTAAATAAAGA 60
DB 1 CCACCGGTCGAAAAACAGGCTGAGAGCAATGTGAGTAAGCAATGTAAATAAAGA 60
QY 61 TTTAAAAATTATTTCTTAATAAAGTACGAATCCCAATGCCAAGGAGCTTACTTCATCTG 120
DB 61 TTTAAAAATTATTTCTTAATAAAGTACGAATCCCAATGCCAAGGAGCTTACTTCATCTG 120
QY 121 CTTCACAAATGATCTTTTCTGATTTCTCAGCCCTGTGGAGACACAGACAGTGGGAA 180
DB 121 CTTCACAAATGATCTTTTCTGATTTCTCAGCCCTGTGGAGACACAGACAGTGGGAA 180
QY 121 CTTCACAAATGATCTTTTCTGATTTCTCAGCCCTGTGGAGACACAGACAGTGGGAA 180
DB 121 CTTCACAAATGATCTTTTCTGATTTCTCAGCCCTGTGGAGACACAGACAGTGGGAA 180
QY 181 AAACCTGGGCTTCTGCTGAGGAGCCACAGAGAACACAGGCTGGAGTGAATTCAGGGCA 240
DB 181 AAACCTGGGCTTCTGCTGAGGAGCCACAGAGAACACAGGCTGGAGTGAATTCAGGGCA 240
QY 181 AAACCTGGGCTTCTGCTGAGGAGCCACAGAGAACACAGGCTGGAGTGAATTCAGGGCA 240
DB 181 AAACCTGGGCTTCTGCTGAGGAGCCACAGAGAACACAGGCTGGAGTGAATTCAGGGCA 240
QY 241 AGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTAAGTGTCTCGGAGTCCCTTTG 300
DB 241 AGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTAAGTGTCTCGGAGTCCCTTTG 300
QY 301 CTGCTCCCCCGTGGAGATCCCTGCGATTTACGAACCCGACGCTCATCGCCCTGGATA 360
DB 301 CTGCTCCCCCGTGGAGATCCCTGCGATTTACGAACCCGACGCTCATCGCCCTGGATA 360
QY 361 ACTTCGGAAGACCACTCTACCTTAATTTGTGCTCAGAACTCAGAGTGGCTGCTCT 420
DB 361 ACTTCGGAAGACCACTCTACCTTAATTTGTGCTCAGAACTCAGAGTGGCTGCTCT 420
QY 421 TAGATCAACACATGCTCAAGGTGATTAACCCGAATTGGAGATGTCAGAAAGCTGCTCT 480
DB 421 TAGATCAACACATGCTCAAGGTGATTAACCCGAATTGGAGATGTCAGAAAGCTGCTCT 480
QY 481 ACCTGAACATCTATGCGCTGCGACAGCGGATACAGGCTCAAGCTCCCGCTTGGTGT 540
DB 481 ACCTGAACATCTATGCGCTGCGACAGCGGATACAGGCTCAAGCTCCCGCTTGGTGT 540
QY 541 GGTTCACAGAGAGTCTTCAAGATGCTGCTCAGCTCCATCTTTATGAGGTGCGCCCTGG 600
DB 541 GGTTCACAGAGAGTCTTCAAGATGCTGCTCAGCTCCATCTTTATGAGGTGCGCCCTGG 600

QY 601 CTGCTATAGAGACGTGCTGTGTGTGTGCTGCTACAGTACCGGCTAGAAATATTGATTTCT 660
DB 601 CTGCTATAGAGACGTGCTGTGTGTGTGCTGCTACAGTACCGGCTAGAAATATTGATTTCT 660
QY 661 TCACCAATGGGATCAGCATGTCTCCGGGAACTGGGCTTTCAAGAACAGGTGGCTGCTC 720
DB 661 TCACCAATGGGATCAGCATGTCTCCGGGAACTGGGCTTTCAAGAACAGGTGGCTGCTC 720
QY 721 TGTCTGGGATCCAGAAAGAACATCGAGTTCTTCGAGGGGACCCGAGCTGTGACATCT 780
DB 721 TGTCTGGGATCCAGAAAGAACATCGAGTTCTTCGAGGGGACCCGAGCTGTGACATCT 780
QY 781 TTGGGAGTCCGCGGAGCCATTAAGTGTTCATGCTTAATACGTCTCCATGGCCAAAG 840
DB 781 TTGGGAGTCCGCGGAGCCATTAAGTGTTCATGCTTAATACGTCTCCATGGCCAAAG 840
QY 841 GCTTATTTCCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTAACCTGGAGGCC 900
DB 841 GCTTATTTCCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTAACCTGGAGGCC 900
QY 901 ATGATTATGAGAGAGTGAAGACCTGAGGTGGTTCATTTCTGTGTAACAATGCGT 960
DB 901 ATGATTATGAGAGAGTGAAGACCTGAGGTGGTTCATTTCTGTGTAACAATGCGT 960
QY 961 CAGACTTGAAGCCCTGCTGAGAGTCCCTGAGGACAAAACCTCCAAAGAGCTGACCC 1020
DB 961 CAGACTTGAAGCCCTGCTGAGAGTCCCTGAGGACAAAACCTCCAAAGAGCTGACCC 1020
QY 1021 TCAGCCAGAAAACAAAGTCTTTCACTCGAGTGTGATGTGTCTTTCTTAATGAGC 1080
DB 1021 TCAGCCAGAAAACAAAGTCTTTCACTCGAGTGTGATGTGTCTTTCTTAATGAGC 1080
QY 1081 CTCTAGATCTATTTGCTGAGAAAGCATTTAAAGCAATTCCTTCATCATGAGGATTA 1140
DB 1081 CTCTAGATCTATTTGCTGAGAAAGCATTTAAAGCAATTCCTTCATCATGAGGATTA 1140
QY 1141 ACCAGAGTGTGGCTTCTGCTGCTGCTTGAAGAGGCTCTGAGATCTCAGTGGCTCA 1200
DB 1141 ACCAGAGTGTGGCTTCTGCTGCTGCTTGAAGAGGCTCTGAGATCTCAGTGGCTCA 1200
QY 1201 ACAAGTCCCTTGGCTCTCATGATTAACAAACATCTGCAATCCCGCTCAGTATTTGC 1260
DB 1201 ACAAGTCCCTTGGCTCTCATGATTAACAAACATCTGCAATCCCGCTCAGTATTTGC 1260
QY 1261 ACCTGTGGCTAATTAATCTTCCATGACAGCAATCCCTGATCTGAATTCGAGACAGTC 1320
DB 1261 ACCTGTGGCTAATTAATCTTCCATGACAGCAATCCCTGATCTGAATTCGAGACAGTC 1320
QY 1321 TTCTGAGCTTGTGAGATGTGTTCTTGTGTGCTCTGCACTGATCAAGCTCGATATC 1380
DB 1321 TTCTGAGCTTGTGAGATGTGTTCTTGTGTGCTCTGCACTGATCAAGCTCGATATC 1380
QY 1381 ACAGAGATGTGTGACCTGTCTATCTTAATGAGTTTCGACCGGCTCAGTCTTTG 1440
DB 1381 ACAGAGATGTGTGACCTGTCTATCTTAATGAGTTTCGACCGGCTCAGTCTTTG 1440
QY 1441 AAGACAGAAACCGGCTTTTGTCAAGCCGACAGCTGATGAATCCGCTTGTGTGTCG 1500
DB 1441 AAGACAGAAACCGGCTTTTGTCAAGCCGACAGCTGATGAATCCGCTTGTGTGTCG 1500
QY 1501 GTGTGCTTCTGAAAGGGGACATTTGTTATGTTGGAAGAGCCAGAGAGAGAGAAAT 1560
DB 1501 GTGTGCTTCTGAAAGGGGACATTTGTTATGTTGGAAGAGCCAGAGAGAGAGAAAT 1560
QY 1561 TACTGAGCCGAAAGATGATGAATATCTGGGCTACCTTTGCTGAAACCGGGAATCTAATG 1620
DB 1561 TACTGAGCCGAAAGATGATGAATATCTGGGCTACCTTTGCTGAAACCGGGAATCTAATG 1620
QY 1621 GGAAGACCTGTCTTGTGGCCAGCTTAATCTGACAGACAGTACTCCAGCTGAGACT 1680
DB 1621 GGAAGACCTGTCTTGTGGCCAGCTTAATCTGACAGACAGTACTCCAGCTGAGACT 1680
QY 1681 TGAACATAGCCTGGAGACAGAGACTCAAAAGAACCGGGGTGGAATTTTGAACAGACCA 1740

Db 841 GCTTATTCACAMAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGCC 900
 Qy 901 ATGATTTATGAGAGAGTGAAGGAGCTGAGAGTGGTTGCACTTTCTGTGTAACTAATGGGT 960
 Db 901 ATGATTTATGAGAGAGTGAAGAGCTGAGAGTGGTTGCACTTTCTGTGTAACTAATGGGT 960
 Qy 961 CAGACTGAGAGCCCTGCTGAGAGTGGTCTGAGAGCAAAACCTCCAGAGAGTGTGACCC 1020
 Db 961 CAGACTGAGAGCCCTGCTGAGAGTGGTCTGAGAGCAAAACCTCCAGAGAGTGTGACCC 1020
 Qy 1021 TCAGCCGAGAGAGAGAGTCTTCACTGAGAGTGGTGTGATGGTCTTTCTTAATGAGC 1080
 Db 1021 TCAGCCGAGAGAGAGAGTCTTCACTGAGAGTGGTGTGATGGTCTTTCTTAATGAGC 1080
 Qy 1081 CTCTAGATCTATGTTGCTGAGAGAGCTTAAAGCAATTCCTTCATCATCGGAGTCAATA 1140
 Db 1081 CTCTAGATCTATGTTGCTGAGAGAGCTTAAAGCAATTCCTTCATCATCGGAGTCAATA 1140
 Qy 1141 ACCAGAGTGGTGGCTTCTGCTGCTGCTATGAGAGAGGCTCTGAGATCTCAGTGGCTCA 1200
 Db 1141 ACCAGAGTGGTGGCTTCTGCTGCTGCTATGAGAGAGGCTCTGAGATCTCAGTGGCTCA 1200
 Qy 1201 ACAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Db 1201 ACAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Qy 1261 ACCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Db 1261 ACCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Qy 1321 TTCTGAGCTTGGTGGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1380
 Db 1321 TTCTGAGCTTGGTGGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1380
 Qy 1381 ACAGAGATGCTGGTGGACCTGTCTATCTTGAATTTGGGACCGGCTCAGTGGCTTGG 1440
 Db 1381 ACAGAGATGCTGGTGGACCTGTCTATCTTGAATTTGGGACCGGCTCAGTGGCTTGG 1440
 Qy 1441 AAGACGAGAGCCGGCTTTGTCAAGCCGACACCGCTGAGAGTCCGCTTGTGGTGG 1500
 Db 1441 AAGACGAGAGCCGGCTTTGTCAAGCCGACACCGCTGAGAGTCCGCTTGTGGTGG 1500
 Qy 1501 GTGGTGGCTTCTGTAAGGGGAGCAATTTGTAATTTGTAAGGAGCCAGAGAGAGAGT 1560
 Db 1501 GTGGTGGCTTCTGTAAGGGGAGCAATTTGTAATTTGTAAGGAGCCAGAGAGAGAGT 1560
 Qy 1561 TACTGAGCCGAGAGATGTAATTAATCTGGGCTACCTTTGCTGAAACCGGAAATCTTAATG 1620
 Db 1561 TACTGAGCCGAGAGATGTAATTAATCTGGGCTACCTTTGCTGAAACCGGAAATCTTAATG 1620
 Qy 1621 GGAACGAGCTGTCTGTGGGCACTTAATCTGAGCTGAGAGTCTCCAGCTGAGCT 1680
 Db 1621 GGAACGAGCTGTCTGTGGGCACTTAATCTGAGCTGAGAGTCTCCAGCTGAGCT 1680
 Qy 1681 TGAACATGAGGCTGAGAGAGAGCTCAAAAGAACCGGGGTGATTTTGGACAGACCA 1740
 Db 1681 TGAACATGAGGCTGAGAGAGAGCTCAAAAGAACCGGGGTGATTTTGGACAGACCA 1740
 Qy 1741 TCCGCTGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 Db 1741 TCCGCTGAGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
 Qy 1801 TCTCTCTCTCTCAGCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1860
 Db 1801 TCTCTCTCTCTCAGCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1860
 Qy 1861 TTTGGTTTCCCTTCTCTCCATTAATTTCTCCGCAATCTTAAGCTTTCTTTCTGAGCTCA 1920
 Db 1861 TTTGGTTTCCCTTCTCTCCATTAATTTCTCCGCAATCTTAAGCTTTCTTTCTGAGCTCA 1920
 Qy 1921 GCTGCTTTCTATGAGGAGTCTTGCAAAACAAGCTGCTTGGCTGATTTATGAGACTTA 1980
 Db 1921 GCTGCTTTCTATGAGGAGTCTTGCAAAACAAGCTGCTTGGCTGATTTATGAGACTTA 1980

Qy 1981 GGAATGATCTTAAGAGATTTCTTTCAACATCAAAAGTGCATTTGCTTGAAGGCA 2040
 Db 1981 GGAATGATCTTAAGAGATTTCTTTCAACATCAAAAGTGCATTTGCTTGAAGGCA 2040
 Qy 2041 CAGATTTCTTCAATTAATTTGGAAGAGGCTGCTTATGAGTGTCAATTAATGCTTT 2100
 Db 2041 CAGATTTCTTCAATTAATTTGGAAGAGGCTGCTTATGAGTGTCAATTAATGCTTT 2100
 Qy 2101 TGTATCTATATGAAATTAATCAAGATGTAATAGGAGAGAGAGAGAGAGAGAGAGAG 2158
 Db 2101 TGTATCTATATGAAATTAATCAAGATGTAATAGGAGAGAGAGAGAGAGAGAGAGAG 2158
 RESULT 4
 US-10-433-256-23
 ; Sequence 23, Application US/10433256
 ; Publication No. US20040081980A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SANJANWALA, Madhusudan M.; YAO, Monique G.
 ; APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
 ; APPLICANT: ARVIZU, Chandra S.; RING, Huijun Z.
 ; APPLICANT: LEE, Ernestine A.; DING, Li
 ; APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
 ; APPLICANT: YUB, Henry; TRIBOULEY, Catherine M.
 ; APPLICANT: LU, Dyrung Aina M.; LAL, Preeti G.
 ; APPLICANT: WARREN, Bridget A.; YANG, Junming
 ; APPLICANT: CHAMLA, Narinder K.; NGUYEN, Daniel B.
 ; APPLICANT: GANDHI, Ameena R.; LU, Yan
 ; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
 ; FILE REFERENCE: PI-0313 USN
 ; CURRENT APPLICATION NUMBER: US/10/433,256
 ; PRIOR FILING DATE: 2003-05-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/47429
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US 60/254,308
 ; PRIOR FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: US 60/256,189
 ; PRIOR FILING DATE: 2000-12-15
 ; PRIOR APPLICATION NUMBER: US 60/257,713
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: US 60/262,706
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: US 60/266,020
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 23
 ; LENGTH: 2232
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incycle ID No: 6538080CB1
 US-10-433-256-23
 Query Match 90.5%; Score 1953; DB 18; Length 2232;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1959; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 176 GGGAAAACTGGGGCTTCTGTGTAAGGGCCACAGAGAAACACAGGGCTGGATGATCA 235
 Db 264 GGTCCAGAGAGGGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGGCTGGATGATCA 323
 Qy 236 GGGCAAGCACTGCTGTGCTGGAAGCCCTGTGCTGTGAACGTGTCTTCGAGAGTCC 295
 Db 324 GGGCAAGCACTGCTGTGCTGGAAGCCCTGTGCTGTGAACGTGTCTTCGAGAGTCC 383
 Qy 296 CTTTGTGCTCCCGCGCTGGGATCCCTGCGATTTAAGAACCCGACGCTGATGCGCTG 355
 Db 384 CTTTGTGCTCCCGCGCTGGGATCCCTGCGATTTAAGAACCCGACGCTGATGCGCTG 443

QY 356 GGATTAACCTTGGAGAACCACTCTCTACCTTAATTTTGGCTCCAGAACTCAGAGTGGCT 415
Db 444 GGATTAACCTTGGAGAACCACTCTCTACCTTAATTTTGGCTCCAGAACTCAGAGTGGCT 503
QY 416 GCTCTTGAATCAACATGCTCAAGGTGCTTACCCGAAATTCGAGTGTCAAGAGACTG 475
Db 504 GCTTTAGATCAACATGCTCAAGGTGCTTACCCGAAATTCGAGTGTCAAGAGACTG 563
QY 476 CTTTACCTTAACATCTATGAGGCTGCGCAAGCCGATACAGGCTCCAGAGTCCCGCTT 535
Db 564 CTTTACCTTAACATCTATGAGGCTGCGCAAGCCGATACAGGCTCCAGAGTCCCGCTT 623
QY 536 GGTGTGTTCCAGAGGTGCTTCAAGACTGCTCAGCTCCATCTTTGATGGTCCG 595
Db 624 GGTGTGTTCCAGAGGTGCTTCAAGACTGCTCAGCTCCATCTTTGATGGTCCG 683
QY 596 CTTGGCTGCTTATGAGAGCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 655
Db 684 CTTGGCTGCTTATGAGAGCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 743
QY 656 TTTCTTACCAACATGGGATCAGCATGCTCCGAGGAACTGGGCTTCAAGAGCAAGGTG 715
Db 744 TTTCTTACCAACATGGGATCAGCATGCTCCGAGGAACTGGGCTTCAAGAGCAAGGTG 803
QY 716 TGTCTGCTGTGCTCCAGAGAAACATCGAGTTCTTGCTGGAGAACCCGAGCTTGTGAC 775
Db 804 TGTCTGCTGTGCTCCAGAGAAACATCGAGTTCTTGCTGGAGAACCCGAGCTTGTGAC 863
QY 776 CATCTTTGGCGAGTCCGCGGAGCCATTAAGTGTCTTACTTATTAATGTTCCCATG 835
Db 864 CATCTTTGGCGAGTCCGCGGAGCCATTAAGTGTCTTACTTATTAATGTTCCCATG 923
QY 836 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGCTGCGCATCATCTTACCTTACCTGA 895
Db 924 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGCTGCGCATCATCTTACCTTACCTGA 983
QY 896 GGGCCATGATTAAGAGAGTGAAGAGCGTGAAGTGTGACATTTCTGTGTGAACA 955
Db 984 GGGCCATGATTAAGAGAGTGAAGAGCGTGAAGTGTGACATTTCTGTGTGAACA 1043
QY 956 TGGCTGAGCTTGAAGGCTGCTGAGTGTCTGAGACAAACCTTCAAGAGAGTCT 1015
Db 1044 TGGCTGAGCTTGAAGGCTGCTGAGTGTCTGAGAGCAAAACCTTCAAGAGAGTCT 1103
QY 1016 GACCTCGCCAGAAACAAAGTCTTTCACTGAGTGTGATGTTCTTTCTTAA 1075
Db 1104 GACCTCGCCAGAAACAAAGTCTTTCACTGAGTGTGATGTTCTTTCTTAA 1163
QY 1076 TGAAGCTCTAGATCTATGCTCAAGAAAGATTTAAAGCAATTCCTTCATCATCGAGT 1135
Db 1164 TGAAGCTCTAGATCTATGCTCAAGAAAGATTTAAAGCAATTCCTTCATCATCGAGT 1223
QY 1136 CAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAGGCTCTGAGATCTCAGTGG 1195
Db 1224 CAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAGGCTCTGAGATCTCAGTGG 1283
QY 1196 CTCCAAAGAGCTTGGCCCTTCATCTGATTAACAACATCCTGACATCCCGCTCAGTA 1255
Db 1284 CTCCAAAGAGCTTGGCCCTTCATCTGATTAACAACATCCTGACATCCCGCTCAGTA 1343
QY 1256 TTTGCACTTGTGCTATGAATCTTCATGACAGACATCCCTGACTGAATCCGAGA 1315
Db 1344 TTTGCACTTGTGCTATGAATCTTCATGACAGACATCCCTGACTGAATCCGAGA 1403
QY 1316 CAGTCTTCTGAGACTTGGAGATGTTCTTTGTGTGTCCCTGCACTGATCAAGCTG 1375
Db 1404 CAGTCTTCTGAGACTTGGAGATGTTCTTTGTGTGTCCCTGCACTGATCAAGCTG 1463
QY 1376 AATATCAGAGATGCTGTGCTCACTCTTATGAGTTCGAGACCGGCTCAGTG 1435
Db 1464 AATATCAGAGATGCTGTGCTCACTCTTATGAGTTCGAGACCGGCTCAGTG 1523
QY 1436 CTTTGAAGACAGAAAGCGGCTTTGTCAAGCCGACCAAGCTGATGAAGTTCGCTTGT 1495

Db 1524 CTTTGAAGACAGAAAGCGGCTTTTGTCAAGCCGACCAAGCTGTATGAAGTCCGTTTGT 1583
QY 1496 GTTCGATGATGCTCTTCTGAAGAGGAGCATTTGTTATGTTCAAGAGACCAAGAGAGA 1555
Db 1584 GTTCGATGATGCTCTTCTGAAGAGGAGCATTTGTTATGTTCAAGAGACCAAGAGAGA 1643
QY 1556 GAAATTAAGAGCGGAGATGATGAATACTGGGCTTACCTTTGCTGAAACCGGAAATCC 1615
Db 1644 GAAATTAAGAGCGGAGATGATGAATACTGGGCTTACCTTTGCTGAAACCGGAAATCC 1703
QY 1616 TAAATGGAACGACCTGTCTGTGTGCGCAGCTTATTAATCTGAATGAGATCTTCAAGT 1675
Db 1704 TAAATGGAACGACCTGTCTGTGTGCGCAGCTTATTAATCTGAATGAGATCTTCAAGT 1763
QY 1676 GGAATTAAGAGAGCTCTGAGACAGAGATCAAGAAACCGGAGTGAATTTTGGACAG 1735
Db 1764 GGAATTAAGAGAGCTCTGAGACAGAGATCAAGAAACCGGAGTGAATTTTGGACAG 1823
QY 1736 CACCATCCCTGATCTGTGTGCTGCTCGACATGCTCAGAGTCTTCTTCTTAAAC 1795
Db 1824 CACCATCCCTGATCTGTGTGCTGCTCGACATGCTCAGAGTCTTCTTCTTAAAC 1883
QY 1796 TTTCTCTCTCTCTGAGGCTTCTTTCTTTTGTGTGCTCTTGAAGATTAATCTTCT 1855
Db 1884 TTTCTCTCTCTCTGAGGCTTCTTTCTTTTGTGTGCTCTTGAAGATTAATCTTCT 1943
QY 1856 GTGATTTTGTGCTTCTTCTTCTTCCATTAATTTCTCCGATCATTAAGTCTTCTGA 1915
Db 1944 GTGATTTTGTGCTTCTTCTTCTTCCATTAATTTCTCCGATCATTAAGTCTTCTGA 2003
QY 1916 GCTCAGCTGCTTCTTATGAGGAGATCTTGAAGAAACAGTCTTCTGATTTATGAG 1975
Db 2004 GCTCAGCTGCTTCTTATGAGGAGATCTTGAAGAAACAGTCTTCTGATTTATGAG 2063
QY 1976 ACTTGAAGATGATCTTACAGATTTCTTTCAACATCAAAAGTGCATTTGTCTGGA 2035
Db 2064 ACTTGAAGATGATCTTACAGATTTCTTTCAACATCAAAAGTGCATTTGTCTGGA 2123
QY 2036 GCGAACAAGTTTCTTCAATTAATTTGGAAGAGGCTGCTATTAAGTGTATTAAT 2095
Db 2124 GCGAACAAGTTTCTTCAATTAATTTGGAAGAGGCTGCTATTAAGTGTATTAAT 2183
QY 2096 GGTTTGTAATCAATTAATTAATTAATCAAGATGTAATTAAGTAATTAAGTAAT 2144
Db 2184 GGTTTGTAATCAATTAATTAATTAATCAAGATGTAATTAAGTAATTAAGTAAT 2232

RESULT 5
US-10-047-249
; Sequence 249, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-047-249

Query Match 88.8%; Score 1917; DB 17; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;
QY 186 GGGCTTCTGCTGAAGGCCACAGAGAAACAGGCTGGAGTGAATTCAGGCCAAGCAA 245

138 GGGGCTTCTGCTGGAAGGGGCAAGAGAAACAGAGCTGGATGATTCAAGGCAAGCAA 197
 246 GTCAATGCTGGGGAAGCCCTGCTGCTGGAACGTGTTCTCTGGAGTCCCTTCTGCT 305
 198 GTCACTGCTGGGGAAGCCCTGCTGCTGGAACGTGTTCTCTGGAGTCCCTTCTGCT 257
 306 CCCCCGCTGGGATCCCTGCGATTTTACGAACCCGAGCCCTGATCCGCTGGATTAATTG 365
 258 CCCCCGCTGGGATCCCTGCGATTTTACGAACCCGAGCCCTGATCCGCTGGATTAATTG 317
 366 CGAAGAGCCACCTCTACCTTAATTT-----GTGCTCCAGAACTCAGAGTGGCT 415
 318 CGAAGAGCCACCTCTACCTTAATTTGTAAGACAGGTGCTCCAGAACTCAGAGTGGCT 377
 416 GCTCTTAATCAACATGCTCAAGTGTATTAACCGAAATTCGGAAGTCAAGAACTG 475
 378 GCTCTTAATCAACATGCTCAAGTGTATTAACCGAAATTCGGAAGTGTCAAGAACTG 437
 476 CCTTACCTGAACATCTATGCGCTGCGCAAGCGATACAGGCTCAAGCTCCCGTCTT 535
 438 CCTTACCTGAACATCTATGCGCTGCGCAAGCGATACAGGCTCAAGCTCCCGTCTT 497
 536 GGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAAGCTCATCTTTGATGGTCCGC 595
 498 GGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAAGCTCATCTTTGATGGTCCGC 557
 536 CCTGCTGCTTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
 558 CCTGCTGCTTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
 656 TTTCTTCAACATAGGAGTCAAGCATGCTCCGGGGAAGTGGGCTTCAAGAACTCAGAGTGGC 715
 618 TTTCTTCAACATAGGAGTCAAGCATGCTCCGGGGAAGTGGGCTTCAAGAACTCAGAGTGGC 677
 716 TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
 678 TGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 776 CATCTTTGGCGAGTCCGCGGAGCCATTAAGTCTTCACTTAATCTGCTCCATGGC 835
 738 CATCTTTGGCGAGTCCGCGGAGCCATTAAGTCTTCACTTAATCTGCTCCATGGC 797
 836 CAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGCCATCATCTTACCTGGA 895
 798 CAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGCCATCATCTTACCTGGA 857
 896 GGCCCATGATTTATGAGAGAGTGAAGACTGCAAGTGTGCACTTTCTGCTGTAACA 955
 858 GGCCCATGATTTATGAGAGAGTGAAGACTGCAAGTGTGCACTTTCTGCTGTAACA 917
 956 TGCGTCAAGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
 918 TGCGTCAAGCTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977
 1016 GACCTTCAGCCAGAAACAAAGTCTTCACTGAGAGTGTGAGTGTGCTTCTTCTTAA 1075
 978 GACCTTCAGCCAGAAACAAAGTCTTCACTGAGAGTGTGAGTGTGCTTCTTCTTAA 1037
 1076 TGAGGCTTGAATCTATTTCTCAGAAAGCAATTTAAGCAATCTTCCATCATGAGT 1135
 1038 TGAGGCTTGAATCTATTTCTCAGAAAGCAATTTAAGCAATCTTCCATCATGAGT 1097
 1136 CAATTAACAAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1195
 1098 CAATTAACAAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
 1136 CTGCAACAGTCCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
 1158 CTGCAACAGTCCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1217
 1256 TTTGCACTTGTGTGCTTAATGATTAATCTTCCATGCAAGCACTCCGATGAATCCGAGA 1315

1218 TTTGCACTTGTGTGCTTAATGATTAATCTTCCATGCAAGCACTCCCTGACTGAATCCGAGA 1277
 1316 CAGCTTCTGGAAGCTTCTTGGAGATGATGTTCTTTGGTCCCTGCACTGATCAGCTCG 1375
 1278 CAGCTTCTGGAAGCTTCTTGGAGATGATGTTCTTTGGTCCCTGCACTGATCAGCTCG 1337
 1376 ATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1435
 1338 ATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1397
 1436 CTTTGAACACAGAACCCGCTTTTGTCAAAAGCCGACCAAGCTGATGAAGTCCGCTTGT 1495
 1398 CTTTGAACACAGAACCCGCTTTTGTCAAAAGCCGACCAAGCTGATGAAGTCCGCTTGT 1457
 1496 GTTGGTGTGCTTCTTGAAGGGGGAATTTATGTTTGAAGGACCAAGGAGAGA 1555
 1458 GTTGGTGTGCTTCTTGAAGGGGGAATTTATGTTTGAAGGACCAAGGAGAGA 1517
 1556 GAAATTAATGAGCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAACCGGGAATCC 1615
 1518 GAAATTAATGAGCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAACCGGGAATCC 1577
 1616 TAATGGAAACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1675
 1578 TAATGGAAACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1637
 1676 GAACTTGAACATGAGCCCTGGAAGAGACTCAAAAGACCGGGGGGATTTTGAACAG 1735
 1638 GAACTTGAACATGAGCCCTGGAAGAGACTCAAAAGACCGGGGGGATTTTGAACAG 1697
 1736 CACCAATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
 1698 CACCAATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757
 1796 TTTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCTGCTGCTGCTGCTGCTGCT 1855
 1758 TTTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCTGCTGCTGCTGCTGCTGCT 1817
 1856 GTGATTTGGTTCCT 1915
 1818 GTGATTTGGTTCCT 1877
 1916 GCTCAGCTGCTTCT 1975
 1878 GCTCAGCTGCTTCT 1937
 1976 ACTTGAATGATCTTCAAGAAATCTTTCACATCAAAAGTGCATTTGCTTGGAA 2035
 1938 ACTTGAATGATCTTCAAGAAATCTTTCACATCAAAAGTGCATTTGCTTGGAA 1997
 2036 GGCAACAAGATTTCTTCAATAATTTGGAAGAGGCTGCTTATAGTTGCTAATAAT 2095
 1998 GGCAACAAGATTTCTTCAATAATTTGGAAGAGGCTGCTTATAGTTGCTAATAAT 2057
 2096 GATTTGTAACTCATATGAATAATCAGAAATG 2130
 2058 GATTTGTAACTCATATGAATAATCAGAAATG 2092

RESULT 6
 US-10-023-515-3
 ; Sequence 3, Application US/10023515
 ; Publication No. US20020182636A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Curtis, Rory A. J.
 ; APPLICANT: Silos-Santiago, Immaculada
 ; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
 ; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
 ; FILE REFERENCE: 10448-122001
 ; CURRENT APPLICATION NUMBER: US/10/023, 515
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 60/256,369
 ; PRIOR FILING DATE: 2000-12-18

```

? PRIOR APPLICATION NUMBER: 60/279,508
? PRIOR FILING DATE: 2001-03-28
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 1746
? TYPE: DNA
? ORGANISM: Homo sapiens
? IS-10-023-515-3

```

Query Match	80.9%	Score 1746	DB 13	Length 1746
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1746	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Oy	96	ATGCAACAGGGA	CTTACTTCA	CTGCTTCA	CATAGTGCTTTT	CTGATTTCA	GCCC	155						
Db	1	ATGCAACAGGGA	CTTACTTCA	CTGCTTCA	CAATGGTGCTTTT	CTGATTTCA	GCCC	60						
Oy	156	CTGTGGGACA	CACAGATGGGG	AAAAA	CTGGG	CGTTCTG	CTGAAGGGCC	ACAAGAAAC	215					
Db	61	CTGTGGGACA	CACAGATGGGG	AAAAA	CTGGG	CGTTCTG	CTGAAGGGCC	ACAAGAAAC	120					
Oy	216	ACCAAGCTGG	ATGATGATTT	CAAGGCA	CAAGTCA	CTGTCTGG	GAAGCCCTG	TCGTG	275					
Db	121	ACCAAGCTGG	ATGATGATTT	CAAGGCA	CAAGTCA	CTGTCTGG	GAAGCCCTG	TCGTG	180					
Oy	276	AAGTGTTCT	CGGAAGT	CCCCCTT	GGCGT	CCCCCG	CTGGGAT	CCCTG	GAATTACGAC	335				
Db	181	AAGTGTTCT	CGGAAGT	CCCCCTT	GGCGT	CCCCCG	CTGGGAT	CCCTG	GAATTACGAC	240				
Oy	336	CCGAGCTG	GCATGCG	CCCTGG	ATTA	CTTGC	GAAGAAC	CACTCTTA	ACCTTAATTGTC	395				
Db	241	CCGAGCTG	GCATGCG	CCCTGG	ATTA	CTTGC	GAAGAAC	CACTCTTA	ACCTTAATTGTC	300				
Oy	396	CTCAGAACT	CAGAGTGG	CTGCA	CTTA	AGATCA	CAATGCT	CAAGTGA	TTATACCGA	455				
Db	301	CTCAGAACT	CAGAGTGG	CTGCA	CTTA	AGATCA	CAATGCT	CAAGTGA	TTATACCGA	360				
Oy	456	TTTGGAGTG	TCAGAGACT	GTGCTT	CACTGA	CACTAT	GTGCG	CTGCG	CCACGCGCAT	515				
Db	361	TTTGGAGTG	TCAGAGACT	GTGCTT	CACTGA	CACTAT	GTGCG	CTGCG	CCACGCGCAT	420				
Oy	516	GGCTCAAG	CTCCCG	CTTGG	TGGTG	TTCC	CAAGAG	TGGT	CTTCA	GTGCTAC	575			
Db	421	GGCTCAAG	CTCCCG	CTTGG	TGGTG	TTCC	CAAGAG	TGGT	CTTCA	GTGCTAC	480			
Oy	576	TCATATTT	GATGGT	CGCG	CTGG	CTGCT	TA	TGAG	AGCAT	GTGCTG	635			
Db	481	TCATATTT	GATGGT	CGCG	CTGG	CTGCT	TA	TGAG	AGCAT	GTGCTG	540			
Oy	636	TACCGGCT	AGGAATAT	TGGT	TTCTT	CAC	CAATGG	ATCA	GCAT	GTCTCC	GGGAACTGG	695		
Db	541	TACCGGCT	AGGAATAT	TGGT	TTCTT	CAC	CAATGG	ATCA	GCAT	GTCTCC	GGGAACTGG	600		
Oy	696	GCCTTCA	AGGAC	CAAGTGG	CTGCT	CTGAT	CTGG	AGTCA	GAAGAA	CATG	GAATTC	755		
Db	601	GCCTTCA	AGGAC	CAAGTGG	CTGCT	CTGAT	CTGG	AGTCA	GAAGAA	CATG	GAATTC	660		
Oy	756	GGGAGCCCC	CACTCTG	TA	CCATCTT	TGG	AGAT	CCG	GGGAG	CCCAT	ATGTTT	CTAGT	815	
Db	661	GGGAGCCCC	CACTCTG	TA	CCATCTT	TGG	AGAT	CCG	GGGAG	CCCAT	ATGTTT	CTAGT	720	
Oy	816	CTTATACT	GTCTCC	CAATGG	CCAAAG	GGCTT	ATTC	CA	CAAG	GCAT	CAT	TGAG	AGTGG	875
Db	721	CTTATACT	GTCTCC	CAATGG	CCAAAG	GGCTT	ATTC	CA	CAAG	GCAT	CAT	TGAG	AGTGG	780
Oy	876	GCATCAT	CTCTTA	CTTGA	AGG	CCCAT	GTAT	TG	AGAA	AGTGA	AGCAT	GTG	AGT	935
Db	781	GCATCAT	CTCTTA	CTTGA	AGG	CCCAT	GTAT	TG	AGAA	AGTGA	AGCAT	GTG	AGT	840
Oy	936	GCACATTT	CTGTG	ATACAA	TGCG	TCAG	ACTCT	TGA	GG	CCCTG	CTGA	GGT	CTGA	995
Db	841	GCACATTT	CTGTG	ATACAA	TGCG	TCAG	ACTCT	TGA	GG	CCCTG	CTGA	GGT	CTGA	900

Qy	996	AAACCTCCAGAGAGCTGTGACCTCAGCCAGAAAACAAAGCTTTCACTCGAGTGTT	1055
Db	901	AAACCTCCAGAGAGCTGTGACCTCAGCCAGAAAACAAAGCTTTCACTCGAGTGTT	960
Qy	1056	GATGTGCTTTCTTTCTTAATGAGCTCTAGATCTAATGTGTCAAGAAAGCATTTAAAGCA	1115
Db	961	GATGTGCTTTCTTTCTTAATGAGCTCTAGATCTAATGTGTCAAGAAAGCATTTAAAGCA	1020
Qy	1116	ATTCTTTCATCATGAGAGTCATAATACACAGAGTGTGCTTCTGTGCTCATGAGAG	1175
Db	1021	ATTCTTTCATCATGAGAGTCATAATACACAGAGTGTGCTTCTGTGCTCATGAGAG	1080
Qy	1176	GCCTCCGAGATCCTGAGTGGCTCCCAAAAGTCCCTTGCCCTTCATCTGATACAAACATC	1235
Db	1081	GCCTCCGAGATCCTGAGTGGCTCCCAAAAGTCCCTTGCCCTTCATCTGATACAAACATC	1140
Qy	1236	CTGCACATCCCGCTCAGTAAATTTGGACCTTGTGTGCTAATGAATACTCCATGACAAGCAC	1295
Db	1141	CTGCACATCCCGCTCAGTAAATTTGGACCTTGTGTGCTAATGAATACTCCATGACAAGCAC	1200
Qy	1296	TCCCTGACCTGAATATCCGAGACAGTCTTCTGGACTTGTGAGATGTGTCTTTGTGTGTC	1355
Db	1201	TCCCTGACCTGAATATCCGAGACAGTCTTCTGGACTTGTGAGATGTGTCTTTGTGTGTC	1260
Qy	1356	CCCTGACCTGATCAGAGCTCGATATCAAGAGATCTGTGTGCACTGTCTAATCTAATGAG	1415
Db	1261	CCCTGACCTGATCAGAGCTCGATATCAAGAGATCTGTGTGCACTGTCTAATCTAATGAG	1320
Qy	1416	TTTGGGACCCGGCTCAGTGTCTTGAAGACAGAAAGCCGGCTTTTGTCAAAAGCCGACAC	1475
Db	1321	TTTGGGACCCGGCTCAGTGTCTTGAAGACAGAAAGCCGGCTTTTGTCAAAAGCCGACAC	1380
Qy	1476	GCTGATGAAGTCCGCTTTGTGTTCGGTGTGACCTTCTGAAAGGGGAGCAATTGTATGTTC	1535
Db	1381	GCTGATGAAGTCCGCTTTGTGTTCGGTGTGACCTTCTGAAAGGGGAGCAATTGTATGTTC	1440
Qy	1536	GAAGGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATAATCTGGGCTACC	1595
Db	1441	GAAGGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATAATCTGGGCTACC	1500
Qy	1596	TTTGTCTGAAACCGGGAAATCTCTAATGGGAAGACCTGTCTCTGTGGCCAGCTTAATAATCTG	1655
Db	1501	TTTGTCTGAAACCGGGAAATCTCTAATGGGAAGACCTGTCTCTGTGGCCAGCTTAATAATCTG	1560
Qy	1656	ACTGAGCAGTACCTCCAGCTGACCTTGAAATGAGAGCTCGGACAGAGACTCAAAAGAACCG	1715
Db	1561	ACTGAGCAGTACCTCCAGCTGACCTTGAAATGAGAGCTCGGACAGAGACTCAAAAGAACCG	1620
Qy	1716	CGGATGGAATTTTGGACACAGACCAATCCCTGTGATCTGTGTGCTCTCCGACATGTCTCAC	1775
Db	1621	CGGATGGAATTTTGGACACAGACCAATCCCTGTGATCTGTGTGCTCTCCGACATGTCTCAC	1680
Qy	1776	AGTCTCTTCTTCTCTTAATCTTTCCCTGTCTCTCCAGAGCTTTCTTTTCTTTTGTGTGCT	1835
Db	1681	AGTCTCTTCTTCTCTTAATCTTTCCCTGTCTCTCTCCAGAGCTTTCTTTTCTTTTGTGTGCT	1740
Qy	1836	CTTTGA 1841	
Db	1741	CTTTGA 1746	

```

RESULT 7
US-10-674-636-3
: Sequence 3, Application US/10674636
: Publication No. US20040086922A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Curtis, Roy A. J.
:
: APPLICANT: Sll0s-Santiago, Inmaculada
:
: TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLTERASE
:
: TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
:
: FILE REFERENCE: 10448-122001
:
: CURRENT APPLICATION NUMBER: US/10/674,636

```

; CURRENT FILING DATE: 2003-09-29
 ; PRIOR APPLICATION NUMBER: US/10/023, 515
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 60/256,369
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: 60/279,508
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1746
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-674-636-3

Query Match 80.9%; Score 1746; DB 18; Length 1746;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 ATGCCACAGGAGCTTACTTCTGCTTCACTGATGCTTTTCTGATTCCTCAGGCC 155
 Db 1 ATGCCACAGGAGCTTACTTCTGCTTCACTGATGCTTTTCTGATTCCTCAGGCC 60
 Qy 156 CTGTTGGACACAGCAGTGGGAAAACTGGCCCTTCTGCTGAAGGCGCAGAGAAC 215
 Db 61 CTGTTGGACACAGCAGTGGGAAAACTGGCCCTTCTGCTGAAGGCGCAGAGAAC 120
 Qy 216 ACCAGGCTGGATGATTCAGGCGCAGAGCTGCTGCTGGAAGCCCTGCTGCTG 275
 Db 121 ACCAGGCTGGATGATTCAGGCGCAGAGCTGCTGCTGGAAGCCCTGCTGCTG 180
 Qy 276 AACGTTCTCGAGTCCCTTTGCTGCTCCCGCTGGAGTCCCTGCTGATTTTGAAC 335
 Db 181 AACGTTCTCGAGTCCCTTTGCTGCTCCCGCTGGAGTCCCTGCTGATTTTGAAC 240
 Qy 336 CCGCAGCTGATCGCCCTGGGATACTTGCAGAGAGCCCTCTCACTTAATTTGTC 395
 Db 241 CCGCAGCTGATCGCCCTGGGATACTTGCAGAGAGCCCTCTCACTTAATTTGTC 300
 Qy 396 CTCCAGAACTCAGATGCTGCTTGAATCAACAGCTCAGAGTGAATACCGGAA 455
 Db 301 CTCCAGAACTCAGATGCTGCTTGAATCAACAGCTCAGAGTGAATACCGGAA 360
 Qy 456 TTGGAGTGAAGAGCTGCTTACCTGAACATCTAGGCTGCTGCTGCTGCTGCTG 515
 Db 361 TTGGAGTGAAGAGCTGCTTACCTGAACATCTAGGCTGCTGCTGCTGCTGCTG 420
 Qy 516 GGTCTCAAGCTCCCGCTTGGTGTGTTCCAGAGAGTGCCTTCAAGCTGCTCAGC 575
 Db 421 GGTCTCAAGCTCCCGCTTGGTGTGTTCCAGAGAGTGCCTTCAAGCTGCTCAGC 480
 Qy 576 TCCATCTTGAATGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 635
 Db 481 TCCATCTTGAATGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Qy 636 TACCGGCTAGGAATTTGTTTCTTCAACAATGGGATCAGAGCTCCGGGGAACG 695
 Db 541 TACCGGCTAGGAATTTGTTTCTTCAACAATGGGATCAGAGCTCCGGGGAACG 600
 Qy 696 GCTTCAAGAGCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
 Db 601 GCTTCAAGAGCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Qy 756 GGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
 Db 661 GGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Qy 816 CTATATCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
 Db 721 CTATATCTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Qy 876 GCATCATCCCTTACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 935

Db 781 GCCATCATCCCTTACTGAGGCCCATGATTAATGAAGAGTGAAGCCTGAGGTGTT 840
 Qy 936 GCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 995
 Db 841 GCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Qy 996 AAACCTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1055
 Db 901 AAACCTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Qy 1056 GATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1115
 Db 961 GATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020
 Qy 1116 ATTCTTCATGATGAGTCAATTAACAGAGTGGCTTCTGCTGCTGCTGCTGCTG 1175
 Db 1021 ATTCTTCATGATGAGTCAATTAACAGAGTGGCTTCTGCTGCTGCTGCTGCTG 1080
 Qy 1176 GCTCCTGAGATCCTGAGTGGCTCAACAAAGTCCCTTGCCTTCATCTGATTAACAATC 1235
 Db 1081 GCTCCTGAGATCCTGAGTGGCTCAACAAAGTCCCTTGCCTTCATCTGATTAACAATC 1140
 Qy 1236 CTGACATCCGCTCAGTATTTGCACTTGTGCTTAATTAATCTTCATGACAAGAC 1295
 Db 1141 CTGACATCCGCTCAGTATTTGCACTTGTGCTTAATTAATCTTCATGACAAGAC 1200
 Qy 1236 TCCCTGATGAAATCCGAGACAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1355
 Db 1201 TCCCTGATGAAATCCGAGACAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Qy 1356 CCTGACATGATCAAGCTGATATCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1415
 Db 1261 CCTGACATGATCAAGCTGATATCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 Qy 1416 TTTGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACAC 1475
 Db 1321 TTTGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACAC 1380
 Qy 1476 GCTGATGAAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1535
 Db 1381 GCTGATGAAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 Qy 1536 GAAGAGCCAGAGAGAGAGAGAGTACTGAGCCGAGAGATGATGAATTAATCTGAGCTACC 1595
 Db 1441 GAAGAGCCAGAGAGAGAGAGAGTACTGAGCCGAGAGATGATGAATTAATCTGAGCTACC 1500
 Qy 1596 TTTGCTGAAACCGGGAATCTTAATGGAAGCACTGCTGCTGCTGCTGCTGCTGCTG 1655
 Db 1501 TTTGCTGAAACCGGGAATCTTAATGGAAGCACTGCTGCTGCTGCTGCTGCTGCTG 1560
 Qy 1656 ACTGAGCAGTACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1715
 Db 1561 ACTGAGCAGTACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 Qy 1716 GGGTGAATTTTGGACAGAGACCAATCCCGTGAATCTGCTGCTGCTGCTGCTGCTGCT 1775
 Db 1621 GGGTGAATTTTGGACAGAGACCAATCCCGTGAATCTGCTGCTGCTGCTGCTGCTGCT 1680
 Qy 1776 AGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1835
 Db 1681 AGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
 Qy 1836 CTTTGA 1841
 Db 1741 CTTTGA 1746

RESULT 8
 US-10-451-168-40
 ; Sequence 40, Application US/10451168
 ; Publication No. US20040091969A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SMITHKLINE BEECHAM CORPORATION

```

; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GPE5039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-451-168-40
```

```

Query Match      76.6%; Score 1652.8; DB 18; Length 1728;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 186 GGGCCCTTCTGTGAAGGGCCACAGAGAAACACAGGCTGGGATGATTCAGGGCAAGCA 245
DB 73 GGGCCCTTCTGTGAAGGGCCACAGAGAAACACAGGCTGGGATGATTCAGGGCAAGCA 132
QY 246 GTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTGTCTCT 305
DB 133 GTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTGTCTCT 192
QY 306 CCCCCGTGGGATCCCTCGATTTACGAACCCGACGCTGCATGCGCTTGGATTAATTG 365
DB 193 CCCCCGTGGGATCCCTCGATTTACGAACCCGACGCTGCATGCGCTTGGATTAATTG 252
QY 366 CGAAGAGCACCTCTTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTATGAT 425
DB 253 CGAAGAGCACCTCTTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTATGAT 312
QY 426 CAACAGATGCTCAAGGTGATTAATCCGAAATTCGAGTGTCAAGAACTGCTTACCTG 485
DB 313 CAACAGATGCTCAAGGTGATTAATCCGAAATTCGAGTGTCAAGAACTGCTTACCTG 372
QY 486 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCAAGCTCCCCGCTTGTGTGTTC 545
DB 373 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCAAGCTCCCCGCTTGTGTGTTC 432
QY 546 CCAGAGAGTGTCTTCAAGACTGTGCTCAGGCTCATCTTTGATGGGTCCGCGCTGAGTGC 605
DB 433 CCAGAGAGTGTCTTCAAGACTGTGCTCAGGCTCATCTTTGATGGGTCCGCGCTGAGTGC 492
QY 606 TATGAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
DB 493 TATGAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 552
QY 666 ACATGGGATCAGCATGTCTCGGGGAACTGGGCTTCAAGGACAGGATGCTGTCTGTGC 725
DB 553 ACATGGGATCAGCATGTCTCGGGGAACTGGGCTTCAAGGACAGGATGCTGTCTGTGC 612
QY 726 TGGGTCCAGAAAGAAATCAGATTCTTCGGTGGGAGCCCAAGCTCTGTGACATCTTTGGC 785
```

```

DB 613 TGGGTCCAGAAAGAAATCAGATTCTTCGGTGGGAGCCCAAGCTCTGTGACATCTTTGGC 672
QY 786 GAGTCCGCGGAGGCCATTAAGTGTTCATGCTTATGCTGTCTCCAGTCCAAAGCTTAA 845
DB 673 GAGTCCGCGGAGGCCATTAAGTGTTCATGCTTATGCTGTCTCCAGTCCAAAGCTTAA 732
QY 846 TTCCCAAAAGCCATGATGAGAGTGGGTGGGCATCATCCCTTACCTGAGAGCCCATAT 905
DB 733 TTCCCAAAAGCCATGATGAGAGTGGGTGGGCATCATCCCTTACCTGAGAGCCCATAT 792
QY 906 TATGAGAAAGTGAAGACCTCGAGGTGTTGCATATTTCTGTGTAAATGCTGTGAC 965
DB 793 TATGAGAAAGTGAAGACCTCGAGGTGTTGCATATTTCTGTGTAAATGCTGTGAC 852
QY 966 TCTGAGGCTCTGTGAGGTGTGCTGAGCAAAACCCCTCCAAAGAGCTGTGACCTCAGC 1025
DB 853 TCTGAGGCTCTGTGAGGTGTGCTGAGCAAAACCCCTCCAAAGAGCTGTGACCTCAGC 912
QY 1026 CAGAAAACAAAGTCTTTCATCTCGAGTGTGATGATGATGATGATGATGATGATGAT 1085
DB 913 CAGAAAACAAAGTCTTTCATCTCGAGTGTGATGATGATGATGATGATGATGATGAT 972
QY 1086 GATCTAATGTCTGAGAAAGCAATTTAAAGCAATTCCTTCCATCATCGAGTCAATTAACAC 1145
DB 973 GATCTAATGTCTGAGAAAGCAATTTAAAGCAATTCCTTCCATCATCGAGTCAATTAACAC 1032
QY 1146 GAGTGTGGCTTCTGTGCTGCTGCTGATGAGAGAGCTCTGAGATCTCGAGTGTGCTCAACAG 1205
DB 1033 GAGTGTGGCTTCTGTGCTGCTGCTGATGAGAGAGCTCTGAGATCTCGAGTGTGCTCAACAG 1092
QY 1206 TCCCTTGGCCCTCCATCTGATTAACAAACATCTGCAATCCGCGCTCAATTTTGACCTT 1265
DB 1093 TCCCTTGGCCCTCCATCTGATTAACAAACATCTGCAATCCGCGCTCAATTTTGACCTT 1152
QY 1266 GTGGCTAATGAATACCTTCCATGACCAAGCACTCCCTGACGAAATCCGAGACAGTCTTTCG 1325
DB 1153 GTGGCTAATGAATACCTTCCATGACCAAGCACTCCCTGACGAAATCCGAGACAGTCTTTCG 1212
QY 1326 GACTGCTTGGAGATATGTTCTTGTGTGCTGCTGCACTGATCAAGCTGATATCAAGA 1385
DB 1213 GACTGCTTGGAGATATGTTCTTGTGTGCTGCTGCACTGATCAAGCTGATATCAAGA 1272
QY 1386 GATGCTGTGACACCTGTCTAATCTTAATGATTTCCGACACCGGCTCAGTGTCTTGAAGAC 1445
DB 1273 GATGCTGTGACACCTGTCTAATCTTAATGATTTCCGACACCGGCTCAGTGTCTTGAAGAC 1332
QY 1446 ACGAAGCCGCTTTTGTCTAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGTGT 1505
DB 1333 ACGAAGCCGCTTTTGTCTAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGTGT 1392
QY 1506 GCTTCTGAAGGGGGGACATTTGTTATGTTCCGAAGAGCCAGAGAGAGAGAAATTACTG 1565
DB 1393 GCTTCTGAAGGGGGGACATTTGTTATGTTCCGAAGAGCCAGAGAGAGAGAAATTACTG 1452
QY 1566 AGCCGGAAGATGATAAATCTGAGCTACCTTGTCTGAAACCGGGAATCTTAATGGGAAC 1625
DB 1453 AGCCGGAAGATGATAAATCTGAGCTACCTTGTCTGAAACCGGGAATCTTAATGGGAAC 1512
QY 1626 GACTGTGCTGTGTGGCCAGCTTAAATCTGACTGACGTAACCTTCAGCTGACTTGAAC 1685
DB 1513 GACTGTGCTGTGTGGCCAGCTTAAATCTGACTGACGTAACCTTCAGCTGACTTGAAC 1572
QY 1686 ATGAGCTTCCGACAGAGACTCAAGAAACCGGAGGTGATTTTGTGAACAGACATATCCC 1745
DB 1573 ATGAGCTTCCGACAGAGACTCAAGAAACCGGAGGTGATTTTGTGAACAGACATATCCC 1632
QY 1746 CTGATCTGTGTGCTCTCGACATGCTCAAGTCTCTTCTTCTTAACTTCTCTCTCT 1805
DB 1633 CTGATCTGTGTGCTCTCGACATGCTCAAGTCTCTCTTCTTCTTAACTTCTCTCTCT 1692
QY 1806 CTGCTCCAGCTTCTTCTTCTTGTGTGCTCTTGA 1841
```

Db 1693 CTCCTCAGACCTTCTTTTCTTTTGTGCTCTGA 1728

RESULT 9
US-10-094-749-736
Sequence 736, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 736
LENGTH: 1962
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-736

Query Match 73.2%; Score 1579.4; DB 17; Length 1962;
Best Local Similarity 92.0%; Pred No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

QY 186 GGGCTTCTGTGAGAGGCGCAGAGGAAACACAGAGCTGGATGATTCAGGGCAAGCA 245
Db 222 GGGCTTCTGTGAGAGGCGCAGAGGAAACACAGAGCTGGATGATTCAGGGCAAGCA 281
QY 246 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTTGCTGCT 305
Db 282 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTTGCTGCT 341
QY 306 CCCCCGCTGGAGATCCCTGGAATTAAGAACCCGAGCCGATGCGCCCGGGATTAAGTTG 365
Db 342 CCCCCGCTGGAGATCCCTGGAATTAAGAACCCGAGCCGATGCGCCCGGGATTAAGTTG 401
QY 366 CGAGAGCCACCTCTTACCTTAATTTGTGCTTCAGAACTCAGATGAGTGTCTTAAAT 425
Db 402 CGAGAGCCACCTCTTACCTTAATTTGTGCTTCAGAACTCAGATGAGTGTCTTAAAT 461
QY 426 CAACATGCTCAGAGTGTATTAACCGAAATTCGAGAGTTCAGAAAGATGCTCTTAAGT 485
Db 462 CAACATGCTCAGAGTGTATTAACCGAAATTCGAGAGTTCAGAAAGATGCTCTTAAGT 521
QY 486 AACATCTATGAGCCGCGCAGCGATACAGGCTCAAGCTCCCGTCTTGTGTTGTTG 545
Db 522 AACATCTATGAGCCGCGCAGCGATACAGGCTCAAGCTCCCGTCTTGTGTTGTTG 581
QY 546 CCAGAGAGTGCCTTCAAGACTGTGCTCAGGCTCCTCATCTTTGATGGGTCCGCTGTGCTC 605
Db 582 CCAGAGAGTGCCTTCAAGACTGTGCTCAGGCTCCTCATCTTTGATGGGTCCGCTGTGCTC 641

QY 606 TATGAGAGCTGTGATTTGTGTCGTCAGTACCGGCTAGGAATATTGTTTCTTAC 665
Db 642 TATGAGAGCTGTGATTTGTGTCGTCAGTACCGGCTAGGAATATTGTTTCTTAC 701
QY 666 ACATGGGATCAGATGCTCCGGGGAATCTGGGCTTCAAGAGCAGGTGCTCTGTGCTC 725
Db 702 ACATGGGATCAGATGCTCCGGGGAATCTGGGCTTCAAGAGCAGGTGCTCTGTGCTC 761
QY 726 TGGGTCCAGAGAAATGAGATTTCTTCGTTGGGAGCCCAAGCTCTGTGACCATTTTGGC 785
Db 762 TGGGTCCAGAGAAATGAGATTTCTTCGTTGGGAGCCCAAGCTCTGTGACCATTTTGGC 821
QY 786 GAGTCCGGGGAGGCATTAAGTTTCTAGTCTTATAGTCTCCCATGGCCAAAGGCTTA 845
Db 822 GAGTCCGGGGAGGCATTAAGTTTCTAGTCTTATAGTCTCCCATGGCCAAAGGCTTA 881
QY 846 TTCCACAAAGCCATCATGAGAGTGGGTGGCCATATCCCTTACCTGAGGCCCATGAT 905
Db 882 TTCCACAAAGCCATCATGAGAGTGGGTGGCCATATCCCTTACCTGAGGCCCATGAT 941
QY 906 TATGAGAGAGTGAAGACCTGAGAGTGTGTCACATTTCTGTGTAAACAATGCTCAGAC 965
Db 942 TATGAGAGAGTGAAGACCTGAGAGTGTGTCACATTTCTGTGTAAACAATGCTCAGAC 1001
QY 966 TCTGAGGCCCTGCTGAGGCTGCTGAGCAAAACCTCCCAAGAGGTGCTGACCCCTGAC 1025
Db 1002 TCTGAGGCCCTGCTGAGGCTGCTGAGCAAAACCTCCCAAGAGGTGCTGACCCCTGAC 1061
QY 1026 CAGAAAACAAAGTCTTTGACCTGAGTGTGATGAGTGTGCTTTCTTCAATAGGCTCTA 1085
Db 1062 CAGAAAACAAAGTCTTTGACCTGAGTGTGATGAGTGTGCTTTCTTCAATAGGCTCTA 1121
QY 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATGTGAGTCAATAACAC 1145
Db 1122 GATCTATTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATGTGAGTCAATAACAC 1181
QY 1146 GAGTGTGCTTCCGCTGCTATGAGAGGCTCTGAGATCTGAGTGGCTCCAAAG 1205
Db 1182 GAGTGTGCTTCCGCTGCTATGAGAGGCTCTGAGATCTGAGTGGCTCCAAAG 1241
QY 1206 TCCCTTGGCCCTCATCTGATCAAAACATCTGACATCCGCTCAGTATTTGACCTT 1265
Db 1242 TCCCTTGGCCCTCATCTGATCAAAACATCTGACATCCGCTCAGTATTTGACCTT 1301
QY 1266 GTGGCTAATGAATCTTCATGACAGCACTCTGACTGAATCCGAGACGTCTTCTG 1325
Db 1302 GTGGCTAATGAATCTTCATGACAGCACTCTGACTGAATCCGAGACGTCTTCTG 1361
QY 1326 GACTTGTGAGAGTGTCTTTGTGTCCTGACATGATCAAGCTCGATATACAGA 1385
Db 1362 GACTTGTGAGAGTGTCTTTGTGTCCTGACATGATCAAGCTCGATATACAGA 1421
QY 1386 GATGCTGTGACCTGTCTACTTCTATGAGTTCCGCAACCGGCTCAAGTCTTTGAAGAC 1445
Db 1422 ----- 1421
QY 1446 ACGAAGCCGCTTTTGTGAAGCCGACACGCTGATGAAGTCGCTTGTGCTGCTG 1505
Db 1422 ----- 1421
QY 1506 GCTTCTGAAAGGGGACATTTGTTATGTTCAAGAGGCGCAGGAGAGGAGATTAAGT 1565
Db 1422 ----- 1421
QY 1566 AGCCGGAAGATGATGAATTAAGTGTGCTACCTTTGCTCAACCGGGAATCTTAATGGAA 1625
Db 1452 AGCCGGAAGATGATGAATTAAGTGTGCTACCTTTGCTCAACCGGGAATCTTAATGGAA 1511
QY 1626 GACTGTCTCTGTGCGACGCTTAATATGACTGAGACGATCTTCAAGTGTGACTTTGAAC 1685
Db 1512 GACTGTCTCTGTGCGACGCTTAATATGACTGAGACGATCTTCAAGTGTGACTTTGAAC 1571

QY 1686 ATGAGCTCGGACAGAGACTCAAGAAACCGGGGGGATTTTGGACACAGACCATGCC 1745
Db 1572 ATGAGCTCGGACAGAGACTCAAGAAACCGGGGGGATTTTGGACACAGACCATGCC 1631
QY 1746 CTGATCTCTGTCTGCTCCGACATGCTCCAGCTCTCTTCTTCTTAACCTTCTCTCT 1805
Db 1632 CTGATCTCTGTCTGCTCCGACATGCTCCAGCTCTCTTCTTCTTAACCTTCTCTCT 1691
QY 1806 CTCTCCAGACCTTCTTCTTCTTCTTGTGCTCTTGAAGATCTTCTGTATTTTGG 1865
Db 1692 CTCTCCAGACCTTCTTCTTCTTGTGCTCTTGAAGATCTTCTGTATTTTGG 1751
QY 1866 TTTCCTCTCTCTCCCAATATTTCTCCGCAATAGCTTCTTCTGAGCTCAGCTGC 1925
Db 1752 TTTCCTCTCTCTCCCAATATTTCTCCGCAATAGCTTCTTCTGAGCTCAGCTGC 1811
QY 1926 TTTCCTAGGGGATCTCTGCAAAACAGCTGCTTCTGTATTTTATGAGCTTAGAGAT 1985
Db 1812 TTTCCTAGGGGATCTCTGCAAAACAGCTGCTTCTGTATTTTATGAGCTTAGAGAT 1871
QY 1986 GATCCTTACGAATTTCTTTCACATCAAAAGTCAATTTGTCTTGAAGGCAAGAGA 2045
Db 1872 GATCCTTACGAATTTCTTTCACATCAAAAGTCAATTTGTCTTGAAGGCAAGAGA 1931
QY 2046 TTTCCTCAATTAATTTGGAAGAGGCTGCGC 2076
Db 1932 TTTCCTCAATTAATTTGGAAGAGGCTGCGC 1962

RESULT 10

US-10-114-270-195

/ Sequence 195, Application US/10114270
/ Publication No. US20040030110A1
/ GENERAL INFORMATION:

/ APPLICANT: Guo, Xiaojia
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Malvankar, Uriel M.
/ APPLICANT: Spytek, Kimberly A.
/ APPLICANT: Patkurajan, Meera
/ APPLICANT: Liu, Ziaohong
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Li, Li
/ APPLICANT: Vernet, Corine
/ APPLICANT: Zerhusen, Bryan D.
/ APPLICANT: Gorman, Linda
/ APPLICANT: Shenoy, Suresh G.
/ APPLICANT: Pena, Carol E.A.
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Burgess, Catherine E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Shinkets, Richard A.
/ APPLICANT: Gangolli, Esna A.
/ APPLICANT: Taupier Jr., Raymond J.
/ APPLICANT: Casman, Stacie J.
/ APPLICANT: Ji, Weizhen
/ APPLICANT: Anderson, David W.
/ APPLICANT: Lieste, Mario W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Edinger, Shlomit R.
/ APPLICANT: Stone, David J.
/ APPLICANT: MacDougall, John R.
/ APPLICANT: Rothenberg, Mark E.
/ TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-322C
/ CURRENT APPLICATION NUMBER: US/10/114, 270
/ PRIOR FILING DATE: 2002-11-27
/ PRIOR APPLICATION NUMBER: 60/281, 086
/ PRIOR FILING DATE: 2001-04-03
/ PRIOR APPLICATION NUMBER: 60/281, 136
/ PRIOR FILING DATE: 2001-04-03
/ PRIOR APPLICATION NUMBER: 60/281, 863

/ PRIOR FILING DATE: 2001-04-05
/ PRIOR APPLICATION NUMBER: 60/281, 906
/ PRIOR FILING DATE: 2001-04-05
/ PRIOR APPLICATION NUMBER: 60/282, 020
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282, 930
/ PRIOR FILING DATE: 2001-04-10
/ PRIOR APPLICATION NUMBER: 60/282, 934
/ PRIOR FILING DATE: 2001-04-10
/ PRIOR APPLICATION NUMBER: 60/283, 512
/ PRIOR FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/283, 710
/ PRIOR FILING DATE: 2001-04-13
/ PRIOR APPLICATION NUMBER: 60/284, 234
/ PRIOR FILING DATE: 2001-04-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 470
/ SEQ ID NO: 195
/ LENGTH: 1746
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1) .. (1744)
/ US-10-114-270-195

Query Match 66.5%; Score 1435.6; DB 18; Length 1746;
Best Local Similarity 95.3%; Pred. No. 0;

Matches 1496; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

QY 176 GGGAAAACTGGGCTTCTGCTGAAGGGCCAGAGAACCAACCAAGCTGGATGATTC 235
Db 69 GGTCCAAAGAGAGGCTTCTGCTGAAGGGCCAGAGAACCAACCAAGCTGGATGATTC 128
QY 236 GGGCAAGCAATCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTCGAGTCC 295
Db 129 GGGCAAGCAATCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTCGAGTCC 188
QY 296 CTTTGTGCTCTCCCGGTGGGATCCCTGCAATTTAGAAACCGGACCTGATCGCTTG 355
Db 189 CTTTGTGCTCTCCCGGTGGGATCCCTGCAATTTAGAAACCGGACCTGATCGCTTG 248
QY 356 GGAATTAATCTGCGAAGACCACTCTTAACCTAATTTGTGCTCGAAGCTCGAGTGGCT 415
Db 249 GGAATTAATCTGCGAAGACCACTCTTAACCTAATTTGTGCTCGAAGCTCGAGTGGCT 308
QY 416 GCTTTAGATCAACATGCTCAAGTGATTAACCGAAATTTGAGTGTGAGAACTG 475
Db 309 GCTTTAGATCAACATGCTCAAGTGATTAACCGAAATTTGAGTGTGAGAACTG 368
QY 476 CCTTACTGTAACATGATGCGCTGCGCAAGCCGATTAAGGCTCCAGCTCCGCTCTT 535
Db 369 CCTTACTGTAACATGATGCGCTGCGCAAGCCGATTAAGGCTCCAGCTCCGCTCTT 428
QY 536 GGTGTGTTCCAGAGGTGCTTCAAGACTGCTCAGCTCATCTTGTGATGGTCCG 595
Db 429 GGTGTGTTCCAGAGGTGCTTCAAGACTGCTCAGCTCATCTTGTGATGGTCCG 488
QY 596 CCGTGTGCTTGAAGAGAGCTGTGTTGTGCTGCTCAATGCGGCTTGAAGAAATTTGG 655
Db 489 CCGTGTGCTTGAAGAGAGCTGTGTTGTGCTGCTCAATGCGGCTTGAAGAAATTTGG 548
QY 656 TTTCTTCAACATGAGATCAGATGCTCCGAGAACTGGGCTTCAAGAGCAAGTGGC 715
Db 549 TTTCTTCAACATGAGATCAGATGCTCCGAGAACTGGGCTTCAAGAGCAAGTGGC 608
QY 716 TGTCTGTCTGTGCTCGAAGAGCATGAGTTCTTGTGTGGAGCCCACTGTGTAC 775
Db 609 TGTCTGTCTGTGCTCGAAGAGCATGAGTTCTTGTGTGGAGCCCACTGTGTAC 668
QY 776 CATCTTGGCGAGTCCGAGGAGGCAATAGTTTCTTATCTGTCTCCATGAC 835
Db 669 CATCTTGGCGAGTCCGAGGAGGCAATAGTTTCTTATGTTTACTGTCTCCATGAC 728

QY 836 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGA 895
DB 729 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGA 788
QY 896 GGGCCCATGATTAAGAGAGTGAAGAGCTGACAGTGGTGGCAATTTCTGTGTGAACA 955
DB 789 GGGCCCATGATTAAGAGAGTGAAGAGCTGACAGTGGTGGCAATTTCTGTGTGAACA 848
QY 956 TGGCTCAGCTTGAGAGCCCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1015
DB 849 TGGCTCAGCTTGAGAGCCCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 908
QY 1016 GACCCCTCAGCAGAAACAAAGCTTTCATCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1075
DB 909 GACCCCTCAGCAGAAACAAAGCTTTCATCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 968
QY 1076 TGAGCCTTGAATCTAATGTTCTCAGAAAGCAATTAAGCAATTCCTTCATCCTGAGT 1135
DB 969 TGAGCCTTGAATCTAATGTTCTCAGAAAGCAATTAAGCAATTCCTTCATCCTGAGT 1028
QY 1136 CATTAACCAAGAGTGGCTTCTGCTGCTTATGAAGAGGCTCCTGAGATCTCAGTGG 1195
DB 1029 CATTAACCAAGAGTGGCTTCTGCTGCTTATGAAGATTTGGCTGCTGCTCATCTGC 1088
QY 1196 CTGCAACAAAGTCCCTGCTGCTCATCTGAATACAA-----ACATCCTGCA 1240
DB 1089 CACTCCCTTAAACCTGATGACGCTTGGCTTCAACGCTGGGCAATTCACAGAAAGCA 1148
QY 1241 CATCCGCTCAGATATTTGACCTGTGGCTAATGATATTCCTCATGACAGCACTCCCT 1300
DB 1149 TATCCGCTCAGATATTTGACCTGTGGCTAATGATATTCCTCATGACAGCACTCCCT 1208
QY 1301 GACTGAATCCGAGACAGTCTTCTGAGCTTGTGAGATATGTTCTTTGTGTGCTCCCTGC 1360
DB 1209 GACTGAATCCGAGACAGTCTTCTGAGCTTGTGAGATATGTTCTTTGTGTGCTCCCTGC 1268
QY 1361 ACTGATCAGAGCTCGATATCAGAGAGTGTGGTGCACCTGCTACTTCTATGAGTTTGC 1420
DB 1269 ACTGATCAGAGCTCGATATCAGAGAGTGTGGTGCACCTGCTACTTCTATGAGTTTGC 1328
QY 1421 GCACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTTGTCAAGCCGACACGCTGA 1480
DB 1329 GCACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTTGTCAAGCCGACACGCTGA 1388
QY 1481 TGAAGTCCGCTTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1540
DB 1389 TGAAGTCCGCTTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1448
QY 1541 AGCCACGAGAGAGAGATTAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1600
DB 1449 AGCCACGAGAGAGAGATTAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1508
QY 1601 TCGAACCGGGAATTCATATGAGAACAGCTGTCTGTGGCAGCTTATATCTGACTGA 1660
DB 1509 TCGAACCGGGAATTCATATGAGAACAGCTGTCTGTGGCAGCTTATATCTGACTGA 1568
QY 1661 GCAATGCTCTGAGCTGAGCTTGAACATGAGCTCGAGCAGACTCAAAAGAACCGGCGGT 1720
DB 1569 GCAATGCTCTGAGCTGAGCTTGAACATGAGCTCGAGCAGACTCAAAAGAACCGGCGAG 1628
QY 1721 GGAATTTTGG 1730
DB 1629 AGATGTGTGG 1638

RESULT 11

US-10-451-168-41
; Sequence 41, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.

APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; PRIOR FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 41
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-41

Query Match 66.4%; Score 1432.4; DB 18; Length 1746;
Best Local Similarity 95.4%; Pred. No. 0;

Matches 1494; Conservative 0; Mismatches 51; Indels 21; Gaps 1;

QY 186 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 245
DB 73 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 132
QY 246 GTCACTGTCTGGAAAGCCCTGTGCTGTGAACGTGTTCTGTGAGATCCCCCTTGTGCT 305
DB 133 GTCACTGTCTGGAAAGCCCTGTGCTGTGAACGTGTTCTGTGAGATCCCCCTTGTGCT 192
QY 306 CCCCCGTGGATCCCTGCGATTTACGAACCCGACGCTGCATGCGCTGGATACTTG 365
DB 193 CCCCCGTGGATCCCTGCGATTTACGAACCCGACGCTGCATGCGCTGGATACTTG 252
QY 366 CGAAGACCACTCTCAACCTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTAGAT 425
DB 253 CGAAGACCACTCTCAACCTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTAGAT 312
QY 426 CAACATCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGCTGCTCTAAGCTG 485
DB 313 CAACATCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGCTGCTCTAAGCTG 372
QY 486 AACATCTAAGGCTCGCCAGCGCATTAAGGCTCAAGCTCCGCTTGTGATGAGTGC 545
DB 373 AACATCTAAGGCTCGCCAGCGCATTAAGGCTCAAGCTCCGCTTGTGATGAGTGC 432
QY 546 CAGAGAGTGCCTTCAAGACTGCTCAGGCTCATCTTGTGATGAGTGCCTGAGTGC 605
DB 433 CAGAGAGTGCCTTCAAGACTGCTCAGGCTCATCTTGTGATGAGTGCCTGAGTGC 492
QY 606 TATGAGAGCTGTGTTGTGTGCTCGTCAATACCGGCTAAGAAATTTGTGTTCTTCAAC 665
DB 493 TATGAGAGCTGTGTTGTGTGCTCGTCAATACCGGCTAAGAAATTTGTGTTCTTCAAC 552
QY 666 ACATGGAGTCAAGATGCTCCGGGGAATGGGCTTCAAGAGACAGATGGCTGTGCTGC 725
DB 553 ACATGGAGTCAAGATGCTCCGGGGAATGGGCTTCAAGAGACAGATGGCTGTGCTGC 612
QY 726 TGGGTTCAGAAAGATCATGAGTTCTTGTGGTGGAGCCCGAGCTCTGTGACATCTTGGC 785

D	613	TGGGTCCAGAAAGAACATCGAGTTCTTGGGTGGGAGCCCAAGCTCTGTGACCATCTTTGGC	672
Q	786	GAGTCCCGGGAGGCCATPAAGTGTTCATGTTCTATACCTGTCTCCATGGCCAAAGGCTTA	845
D	673	GAGTCCCGGGAGGCCATPAAGTGTTCATGTTCTATACCTGTCTCCATGGCCAAAGGCTTA	732
Q	846	TTCCAAAGGCAATCATGAGAGAGTGGGGTGGCATACTCCCTTAACCTGGAGGCCATGAT	905
D	733	TTCCAAAGGCAATCATGAGAGAGTGGGGTGGCATACTCCCTTAACCTGGAGGCCATGAT	792
Q	906	TATGAGAAAGTGAAGACTTGCAGGTGGTTGCAATTTCTGTGTAAACAATCGTCAAGAC	965
D	793	TATGAGAAAGTGAAGACTTGCAGGTGGTTGCAATTTCTGTGTAAACAATCGTCAAGAC	852
Q	966	TTGAGAGCCCTGGCGAGGTGCGTGAAGACAAAACCTCCAAAGAGCTGTTACCTCAAGC	1025
D	853	TTGAGAGCCCTGGCGAGGTGCGTGAAGACAAAACCTCCAAAGAGCTGTTACCTCAAGC	912
Q	1026	CAGAAAACAAAGTCTTCACTCGAGTGGTGAATGTGTGCTTCTTCTTAATAGGCTCTTA	1089
D	913	CAGAAAACAAAGTCTTCACTCGAGTGGTGAATGTGTGCTTCTTCTTAATAGGCTCTTA	972
Q	1086	GATCTATTGTCTCAGAAAGATTTAAGCAATTCCTTCCATCATGAGAGTCATTAACAC	1145
D	973	GATCTATTGTCTCAGAAAGATTTAAGCAATTCCTTCCATCATGAGAGTCATTAACAC	1032
Q	1146	GAGTGTGGCTTCTGTGCTTGAAGAGAGCTCTCGAGATCCTCAGTGGCTCAACAG	1205
D	1033	GAGTGTGGCTTCTGTGCTTGAAGAGAGCTCTCGAGATCCTCAGTGGCTCAACAG	1092
Q	1206	TCCCTT-----GCCCTCAATCGATACAAACATCTGCAATC	1244
D	1093	AACCGTAGTCAGACTTTGGCTTCAACAGTGGGCAATTCACAGAGACATCAGACATC	1152
Q	1245	CCGCTCAGTATTTCGACCTTGTGGCTAATGATACTTCATCAAGACATCCCTGACT	1304
D	1153	CCGCTCAGTATTTCGACCTTGTGGCTAATGATACTTCATCAAGACATCCCTGACT	1212
Q	1305	GAAATCCGAGACAGTCTTGTGACTTGGAGATGTGTTCTTTGTGTGCTCCCTGCACTG	1364
D	1213	GAAATCCGAGACAGTCTTGTGACTTGGAGATGTGTTCTTTGTGTGCTCCCTGCACTG	1272
Q	1365	ATCAGAGCTCGATATCAGAGAGTCTGTGCACTGTCTACTCTTAATGATTTGGCAC	1424
D	1273	ATCAGAGCTCGATATCAGAGAGTCTGTGCACTGTCTACTCTTAATGATTTGGCAC	1332
Q	1425	CGGCTCAGTGGTTTGAAGACAGAGCGGCTTTGTCAAGCGGACACAGCTGATGAA	1484
D	1333	CGGCTCAGTGGTTTGAAGACAGAGCGGCTTTGTCAAGCGGACACAGCTGATGAA	1392
Q	1485	GTCGGCTTGTGTGATGATGTCCTTCTGTAAGGGGAGCAATGTATGTTCGAAGAGCC	1544
D	1393	GTCGGCTTGTGTGATGATGTCCTTCTGTAAGGGGAGCAATGTATGTTCGAAGAGCC	1452
Q	1545	ACGGAGGAGGAAAGTTACTGAAGCCGGAAGATGATGAAATACTGGGCTACCTTGTCTGA	1604
D	1453	ACGGAGGAGGAAAGTTACTGAAGCCGGAAGATGATGAAATACTGGGCTACCTTGTCTGA	1512
Q	1605	ACGGGAAATCTTAATGGGAACGACTGTCTCTGTGGCAGACTTAATATGACTGAGCAG	1664
D	1513	ACGGGAAATCTTAATGGGAACGACTGTCTCTGTGGCAGACTTAATATGACTGAGCAG	1572
Q	1665	TACCTCAGCTGAGACTTGAACATGAAGCTCGGACAGAGATCTCAAGAACCCGGGGTGAT	1724
D	1573	TACCTCAGCTGAGACTTGAACATGAAGCTCGGACAGAGATCTCAAGAACCCGGGGTGAT	1632
Q	1725	TTTTGG 1730	
D	1633	GTGTGG 1638	

```

; Sequence 20, Application US/10381898
; Publication No. US20040086887A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valda; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKY, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIFIN, Jennifer A.;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; LAL, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dyung Anna M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANMALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junning; YAO, Yuming; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-023 USN
; CURRENT APPLICATION NUMBER: US/10-381,898
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ. ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7473645CB1
; US-10-381-898-20

```

RESULT 12
US-10-381-898-20

	Query Match Best Local Similarity	64.2%; 94.7%	Score 1386.2;	DB 18;	Length 1857;
	Matches 1473;	Conservative 0;	Mismatches 13;	Indels 69;	Gaps 1
QY	176 GGGAAAACCTGSGCCCTTCTGTGAAGGCCACAAGAAGAACACCAGCGCTGGATGATTCA				235
Db	264 GGTCACAGAGGGSCCTTTGTCTGTAAGGGCCACAAGAAGAACACAGCGCTGGATGATTCA				323
QY	236 GGGCAAGCAGCATCTGTGCTGGGAAGCCCTGAGCCCTGTGAACCTGTCTTCCTGGAGCTCC				295
Db	324 GGGCAAGCAGCATCTGTGCTGGGAAGCCCTGTGCTGTGAACGATCTTCCTGGAGCTCC				383
QY	296 CTTTGCTGTCCCCCGCTGGGATCTCTGCATTTACGAACCCGCAAGCTGCATTCGCTCG				355
Db	384 CTTTGCTGTCCCCCGCTGGGATCTCTGCATTTACGAACCCGCAAGCTGCATTCGCTCG				443
QY	356 GGATAACTTGGAGAAAGCACCTCTCAACCTTAATTTGTGCTCTCCGAAGACTCGAAGTGGCT				415
Db	444 GGATAACTTGGAGAAAGCACCTCTCAACCTTAATTTGTGCTCTCCGAAGACTCGAAGTGGCT				503
QY	416 GGTCTTAATTAACAATGCTCAAGGTGCATTACCGGAATTTGGAGTGTGAGAAAGCTG				475
Db	504 GGTCTTAATTAACAATVGTCTCAAGGTGCATTACCGGAATTTGGAGTGTGAGAAAGCTG				563
QY	476 CCTTACCTGMAACATCATATGCGCTGCCACGCGCATACAGGCTCTCAAGCTCCCCGATTTT				535


```
Db 564 CCTTACCTGAAACATCTATGCGCTGCGCAAGCCGATACAGGCTCAAGCTCCCGTCT 623
Qy 536 GGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAAGGTCCG 595
Db 624 GGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAAGGTCCG 683
Qy 596 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
Db 684 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 743
Qy 656 TTTCTTCAACAATGAGATCAGCATGCTCCGGGAACTGGGCTTCAAGACAGGTG 715
Db 744 TTTCTTCAACAATGAGATCAGCATGCTCCGGGAACTGGGCTTCAAGACAGGTG 803
Qy 716 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
Db 804 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
Qy 776 CATCTTGGCGAGTCCGGGGAGCCATAGTCTTCTAGTCTTACTGCTCCCATGG 835
Db 864 CATCTTGGCGAGTCCGGGGAGCCATAGTCTTCTAGTCTTACTGCTCCCATGG 923
Qy 836 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACCTG 895
Db 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACCTG 983
Qy 896 GGGCCATATTTAGAGAGAGTGAAGACTGAGAGTGGTGTGCAATTTCTGTGTACA 955
Db 984 GGGCCATATTTAGAGAGAGTGAAGACTGAGAGTGGTGTGCAATTTCTGTGTACA 1043
Qy 956 TGGGTCAACTCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 1015
Db 1044 TGGGTCAACTCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 1103
Qy 1016 GACCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1075
Db 1104 GACCTCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1163
Qy 1076 TGAAGCTTCAATCTATGCTCAGAGAGCTTAAAGCAATCTTCCATCATGAGT 1135
Db 1164 TGAAGCTTCAATCTATGCTCAGAGAGCTTAAAGCAATCTTCCATCATGAGT 1223
Qy 1136 CAATTAACAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1195
Db 1224 CAATTAACAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
Qy 1196 CTCCAAAGTCCCTGCTCATCTGATACAAAACATCTGCAATCCGCTCAGTA 1255
Db 1256 -----TGCACTCCGCTCAGTA 1274
Qy 1256 TTTGACCTTGTGCTAATGAATCTTCATGACAGACCTCCCTGACTGAATCCGAGA 1315
Db 1275 TTTGACCTTGTGCTAATGAATCTTCATGACAGACCTCCCTGACTGAATCCGAGA 1334
Qy 1316 CAGCTTCTGAGCTTGTGAGAGTGTCTTGTGAGTGTGCTGAGTGTGCTGAGTGTG 1375
Db 1335 CAGCTTCTGAGCTTGTGAGAGTGTCTTGTGAGTGTGCTGAGTGTGCTGAGTGTG 1394
Qy 1376 ATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1435
Db 1395 ATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1454
Qy 1436 CTTTGAAGACAGAGAGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGT 1495
Db 1455 CTTTGAAGACAGAGAGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGT 1514
Qy 1496 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1555
Db 1515 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1574
Qy 1556 GAAATTAAGAGCTGAGAGATGATGAATCTGGGCTACCTTGTGCTGCAAGCGGGAATCC 1615
```

```
Db 1575 GAAATTAAGAGCTGAGAGATGATGAATCTGGGCTACCTTGTGCTGCAAGCGGGAATCC 1634
Qy 1616 TAAATGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1675
Db 1635 TAAATGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1694
Qy 1676 GGAATTTGAAGAGCTGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1730
Db 1695 GGAATTTGAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1749

RESULT 13
US-10-233-933A-3
; Sequence 3, Application US/10233933A
; Publication No. US20040214171A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Tetsuro
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG FP02US006
; CURRENT FILING DATE: 2002-09-03
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Felis catus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1803)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1804)..(2145)
; OTHER INFORMATION: n is a, c, g, or t
US-10-233-933A-3

Query Match 53.6%; Score 1156; DB 20; Length 2145;
Best Local Similarity 78.2%; Pred. No. 0;
Matches 146; Conservative 0; Mismatches 386; Indels 22; Gaps 6;

Qy 186 GGGCTTGTGCTGAGAGGCGACAGAGAAACACAGAGCTGGAGTTCAGGGCAAGCA 245
Db 238 GGGCAAGTGTGCTGAGTGCACAGTGAAGAGACAGAGCTGGAGTGGTCCGGGGAAGCA 297
Qy 246 GTCACTGTGCTGAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
Db 298 ACCACTGACTGAGAGAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
Qy 306 CCCCCGTGGAGTCCCTGCGATTTACGAACCCGACGCTGCATCCCTCGGATTAATCTTG 365
Db 358 CTTCTCTAGGCGCCCTGCGATTTAAGCAACCAAGCCCTGCTGCTCCGGGAATGACTTC 417
Qy 366 CGAGAGCACTCTTACCTTAATTTGCTCCGAACTGAGAGTGGCTGCTTAAGT 425
Db 418 GAAATGCAACCTTACCTTAATTTGCTCCGAACTGAGAGTGGCTGCTTAAGT 477
Qy 426 CAACATGCTCAAGGTCATTAACCGAAATTCGAGTGTGCAAGAGTGCCTTAACCTG 485
Db 478 CAACATGCTCAAGGTCATTAACCGAAATTCGAGTGTGCAAGAGTGCCTTAACCTG 537
Qy 486 AACATCTATGCGCTGCGCAAGCGCATACAGGCTCMAAGCTCCCGTCTTGTGTGCTTC 545
Db 538 AACATCTATGCGCTGCGCAAGCGCATACAGGCTCMAAGCTCCCGTCTTGTGTGCTTC 597
Qy 546 CCAGAGAGCTCTTAAGAGTGGCTCAGCTCATCTTGAATGGGTCCGCTGCTGCTGCTGCT 605
Db 598 CCGGGGGGCTCTCAAGAGTGGCTCAGCTCATCTTGAATGGGTCCGCTGCTGCTGCTGCT 657
Qy 606 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
```

Dh 658 TACGAGACGTGCTGATGCTGACTACCCAGTACCGGCTAGAGAAATTTGGTTTTTTCGAC 717
Qy 666 ACATGGGATCGAGTGTCTCCGGGAACTGGGCTTTCAAGACCAAGTGGCTGTCTTC 725
Dh 718 ACAGGGGATGAGTATGCCCCGGGAACTGGGCTTTGGCTGACACAGTGGCTGCTCAAC 777
Qy 726 TGGGTCCAGAAACATGAGTCTTGGTGGGAGCCCGAGCTGTGACATCTTTGGC 785
Dh 778 TGGGTCCGGGACATCGAGTCTTGGTGGTGAACCAAGCTTCGACCACTTTTGA 837
Qy 786 GAGTCCCCGGGAGGACATGAGTCTTGAATCTTACTGTCTCCCATGGCCAAAGCTTA 845
Dh 838 GAGTCAACGGGAGGACATGAGTCTTGAATCTTACTGTCTCCCATGACCAATGGCTTA 897
Qy 846 TTCCACAAGACCATCATGAGAGTGGGGTGGCAATCCCTTAACCTGGAG--GCCCAT 902
Dh 898 TTCCACAAGACCATCATGAGAGTGGGGTGGCAATCCCTTAACCTGATGAGAGACCCCT 957
Qy 903 GATTATGAGAAAGTGAAGACTTGCAGTGGTTGCAATTTCTGTGTAAACAATGGCTGA 962
Dh 958 GGTGATGAGAGAAAGATTTGCAAGTGTCTTGGCGTATCTGTGGTGGCCATGGCTCT 1017
Qy 963 GACTGTAGGCCCCGTGAGTGGCTGAGGACCAAAACCTTCCAGAGAGCTGTGACCTTC 1022
Dh 1018 GACTGTGCTGCCCCGTGAGTGGCTGAGGACCAAAACCTTCCAGAGAGTGTGACATC 1077
Qy 1023 AGCCAGAAAACAAAGTCTTCACTCGAGTGTGATGATGTGCTTCTTCTTAATGAGCCT 1082
Dh 1078 AGCAAGAAATCACTGTTTCCATTCAGATGATGATGATGATGATGATGATGATGATG 1137
Qy 1083 CTAGATCTTATGTCTCAAGAAAGATTTAAAGCAATCTTCCATCATCGAGTCAATAAC 1142
Dh 1138 GTAGCCCTTATGATCAAAAGATTTAAATTCAGTCTTCTATCATCGAGTCAATAAC 1197
Qy 1143 CAGAGTGTGGCTCTGCTGCTGCTATGAGAGAGGCTCCGAGATCTCATGAGTCTCAAC 1202
Dh 1198 CAGAGTGTGGCTCTGCTGCTGCTATGAGAGAGGCTCCGAGATCTCATGAGTCTCAAC 1254
Qy 1203 AGTCCCTTGGCTCTGCTGCTGCTATGAGAGAGGCTCCGAGATCTCATGAGTCTCAAC 1262
Dh 1255 AGTCCCTTGGCTCTGCTGCTGCTATGAGAGAGGCTCCGAGATCTCATGAGTCTCAAC 1314
Qy 1263 CTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
Dh 1315 CTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1374
Qy 1323 CTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1382
Dh 1375 CTGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1434
Qy 1383 AGAGATGCTGAGTCACTGTCTATCTTATGAGTTCGGACCGGCTCAAGTCTTTGAA 1442
Dh 1435 AGAGATGCTGAGTCACTGTCTATCTTATGAGTTCGGACCGGCTCAAGTCTTTAAC 1494
Qy 1443 GACAGAGAGCGGCTTTTGTCAAAAGCCGACCGCTGATGAGTCCGCTTGTGTGGT 1502
Dh 1495 GACAGAGAGCGGCTTTTGTCAAAAGCCGACCGCTGATGAGTCCGCTTGTGTGGT 1554
Qy 1503 GGTGCTTCTTCAAGAGGAGGACATTTGATGTTCAAGAGGCGACGAGAGAGAGATTA 1562
Dh 1555 GGTGCTTCTTCAAGAGGAGGACATTTGATGTTCAAGAGGCGACGAGAGAGAGATTA 1614
Qy 1563 CTGAGAGCGGAGATGATGAAATACCTGGGCTACCTTTGCTGCAACGGGAAATCTTAATGG 1622
Dh 1615 CTGAGAGCGGAGATGATGAAATACCTGGGCTACCTTTGCTGCAACGGGAAATCTTAATGG 1674
Qy 1623 AACGACCTGTCTGTGGGCGACCTTATATCTGAATGAGAGTACCTCAAGTGAATTTG 1682
Dh 1675 GAGGTGTGCTGTGTGGGCGACCTTATATCTGAATGAGAGTACCTCAAGTGAATTTG 1734
Qy 1683 AACATGAGCTCTGAGACAGATCTAAAGAACCGGGGTGATTTTGGACCAAGACCAT- 1741
Dh 1735 AGTGTGAGCTGGGACAGAACTGAAGAGACAGAGGTGAGTTTGGATGAATACCAT 1794

Qy 1742 -CCCCGTGATCTGTCTGCTGCTCCGACATGCTCCAGATGCTCTTTCTTCTTAATTTCC 1800
Dh 1795 GTCCCTGATATACCCCACTCCAGAGGCTTCCAGATGCTCTTCTTCTTAATTTCC 1854
Qy 1801 TCTCTCTCTCCAGGCTTTCTTTTCTTTTGTGCTCTTGAAGATTAATCTTGTGAT 1860
Dh 1855 TTCTTTGTCTCCGCTGCTGCTTNTTNTTCTGCTCAATGGGAAGTCTCTGATGAT 1914
Qy 1861 TTTGGTTTCCCTTCTCTCCCAATATTTCTCCCGAATCATTTACTTTCTTCAAGCTCA 1920
Dh 1915 TTGTTTCTCTT-----TTCCAAATTTTACCCGCGTCTTACTCTCGTT-----T 1962
Qy 1921 GCTGCTTCTATGGGATCTTGCAGAAACAGCTGCTTCCGCTGATATTTATGACTTA 1980
Dh 1963 CTGTGTATGATTAAGTACTCTTGTGAAATGATGATGATGATGATGATGATGATG 2022
Qy 1981 GGAATGATCTTACAG--AATTTCTTCAACATCAAAAGTCAATTTGTCTTGAAGGC 2038
Dh 2023 AGGGATTTTGAAGGAAATTTCTTTCACACCAAAACAATGCTACCGGCTTGAAGGC 2082
Qy 2039 AACAAATTTCTTC 2052
Dh 2083 TACCGATCTTTC 2096

RESULT 14
US-10-233-933A-1
; Sequence 1, Application US/10233933A
; Publication No. US2004024171A1
; GENERAL INFORMATION:
; APPLICANT: Miyazaki, Teisuro
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG FP02US006
; CURRENT APPLICATION NUMBER: US/10/233,933A
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: JP2002-057908
; PRIORITY FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Felis catus
US-10-233-933A-1

Query Match 49.3%; Score 1063.2; DB 20; Length 1629;
Best Local Similarity 80.9%; Pred. No. 4e-296;
Matches 1265; Conservative 0; Mismatches 293; Indels 6; Gaps 2;

Qy 186 GGGCTTGTGCTGAGAGGCGCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGCA 245
Dh 64 GGGCGAGCTGTGATGACACAGTGAAGACACAGGCTGGATGATTCAGGGCAAGCA 123
Qy 246 GTCACTGTCTGGAAGCCCTGTGCTGTAAGTGTCTCTCGAGTCCCTTGTGCTGT 305
Dh 124 ACCACTGTACTGGGAAGCACCGTGTGTGAAACATGTTCTCTGGAATCCCTATGCTCA 183
Qy 306 CCCCCGCTGGGATCCCTGTGATTTACGAACCGGACGCTGATGCGCTGGGATTAATCTG 365
Dh 184 CTTCTCTAGAGGCGCTGTGATTTAAGCAACCAAGCTGTCTGCGGGAAATGACTTC 243
Qy 366 CGAGAGACCACTCTTACCTTAATTTGTGCTCCGAGACTGAGAGTGTGCTGTAGAT 425
Dh 244 GGAATGCAATCTTACCTTAATTTGTGCTCCGAGACTGAGAGTGTGCTGTAGAT 303
Qy 426 CAACATGCTCAAGTGTGATTCAGAAATTCGAGTGTGAGAAAGTGTGCTGTACTG 485
Dh 304 CAACAGTGTCTCAAGTGTGATTCAGAAATTCGAGTGTGAGAAAGTGTGCTGTACTG 363
Qy 486 AACATCTATGGGCTGCGCAAGCGGATACAGGCTTCAAGTCTCCGCTTGTGTGCTTC 545

```

Db 364 AACATCTATGCGCCAGCCCATGCGGACATGCGCTCCAACTCTGTCATGCTGTTTC 423
Qy 546 CCAGAGGAGGCTTGAAGCTGCTCAGCTCCTCATCTTTGATGGGCGCCCTGGCTGCC 605
Db 424 CCGGGGGGCTCTCAAGATGAGCTCAGCTTCTCTCTCTCTGATGGGCTCGCCCTGGCTGCC 483
Qy 606 TATGAGAGCGTGTGCTGTGTGCTGCTCAATACCGGCTTGAAGAAATTTGGTTTCTTCAAC 665
Db 484 TACGAGAGCGTGTGCTGATGCTACCCGCTACCGGCTTGAAGAAATTTGGTTTCTTCAAC 543
Qy 666 ACATGGGATACAGATGCTCTCGGGGAACTGGGCTTCAAGAGACAGAGTGGCTCTGTCTC 725
Db 544 ACAGGGGATGAGACATGCGCGGGGAACTGGGCTTGTGGACAGAGTGGCTCTGTCTC 603
Qy 726 TGGGTCCAGAAAGATCGAGTCTTCCGTTGGGGGACCCGAGCTCTGTGACATCTTTGGC 785
Db 604 TGGGTCCGGGACACATGAGATTTCTTCGGTGGTACCCAGCTCTGTGACATCTTTGGA 663
Qy 786 GAGTCCGGGGGAGCCATGATGCTTCTAGTCTTATATGCTCTCCATGGCCAAAGGCTTA 845
Db 664 GAGTCAAGGGGAGCCATGATGCTTCCAGCTCATTTCTGTCTCCCATAGCCAAATGGCTTA 723
Qy 846 TTCCACAAAGCCATATGAGAGTGGGGTGGCCATCATCTCTTACCTTGAG--GCCAT 902
Db 724 TTCCACAAAGCCATATGAGAGTGGGGTGGCCATCTGCTCTTATCTGATGAGACCCCT 783
Qy 903 GATTATGAAAGATGAGGACCTGAGAGTGTGACATTTCTGTGTGAACAATGGCTTA 962
Db 784 GGTATGAGAGAGAGAGATTTGACGCTGGCTGGCGATCTGTGTGGCCATGGCTT 843
Qy 963 GACTCTGAGGCGCTCTGAGTGTGCTGAGACAAACCTCCAAAGAGCTGCTGACCTC 1022
Db 844 GACTCTGCGCCCTCTGAGTGTGCTGAGGAGGAAACCTCCGAGAGATGATGACATC 903
Qy 1023 AGCCAGAAACAAAGCTTTTCACTGAGTGTGTGATGTGCTTTCTTCTAATAGCTT 1082
Db 904 AGCAAGAAATCTACCTTTTCCATTCAGATGATGATGATTTCTTCTCTGATAGCCT 963
Qy 1083 CTAGATCTATGTCTCAGAAAGCATTTAAAGAAATTCCTTCATCATGGAGTCAATAC 1142
Db 964 GTAGCCCTATTTGACTCAAAAGCATTTAATTAAGTTCTTCTATCATGAGATCAATAC 1023
Qy 1143 CACGAGTGTGCTCTGCTGCTTGAAGAGGCTCTGAGATCTGAGTGGCTCAAC 1202
Db 1024 CACGAGTGTGCTCTCTCTGCTGCTGCTGAGTGTGCTGAGATCTGAGTGGCTCAAC 1080
Qy 1203 AAGTCTGCTGCTCTCATCTGATCAAAACATCTGCAATCCGCTCAATGATTTGCAC 1262
Db 1081 AAGTCTGCTGCTCTCATCTGATCAACAGTTCCTGAATATTCACCAACCATGATTTGCAC 1140
Qy 1263 CTGTGTGCTAATGATTAATCTTCATGACAGCACTCCCTGACTGAATCCGAGACGCTT 1322
Db 1141 CTGTGTGCTGATCAATTAATCTTCAACAAAGCACTCCCGGTGAATGAGATGATTT 1200
Qy 1323 CTGAGCTGTGAGATGATGTTCTTTGTGATCCCTGCACTGATCAAGCTGATATC 1382
Db 1201 CTGAGCTGTGAGATGATGCTCTTTGTGATCCCTGAGGATGATCAAGCTGATATC 1260
Qy 1383 AGAGATGCTGTGCACTGTCTACTTCTAATGATTTGGGACCGGCTCAAGTCTTTGAA 1442
Db 1261 AGAGATGCTGTGCACTGTCTACTTCTAATGATTTGCAACCCGCTCAAGTCTTTAAC 1320
Qy 1443 GACAGGAGCGGCTTTTGTCAAAAGCCGACAGCTGATGAAAGTCCGCTTTGTGGT 1502
Db 1321 GACAGGAGCGGCTTTTGTGAAAGCCGATCACTGATGAATCCGCTTTGTGGT 1380
Qy 1503 GGTGCTTCTGTAAGGGGACATTTGTATGTTGTAAGAGCCAGGAGGAGAGATTA 1562
Db 1381 GGTGCTTCTGTAAGGGGACATTTGTATGTTGTAAGAGCCAGGAGGAGAGATTA 1440
Qy 1563 CTGAGCCCGAGAGATGATGAAATATCTGGGCTACCTTTGCTCGAACCGGAGATCTTA 1622
Db 1441 CTGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500

```

```

Qy 1623 AACGACTGTCTCTGTGGCCAGCTTATATCTGACTGAGCAGTACCTCAGCTGACTTG 1682
Db 1501 GAGGTGTGCTCTGTGGCCAGCTTACACCCAGAGCGAGCAGTACCTGAGCTGAGTTTG 1560
Qy 1683 AACATGAGCTCTGACAGAGACTCAAGAAACCGCGGTGATTTTGGACCAACCATC 1742
Db 1561 AGTGTGAGCGTGGGACAGAAACTGAAGAGCAGAGTGGAGTTTGTGATGATATCAAT 1620
Qy 1743 CCCC 1746
Db 1621 GTCC 1624

RESULT 15
US-10-451-168-42
; Sequence 42, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-42

Query Match 47.3%; Score 1021.4; DB 18; Length 1071;
Best Local Similarity 99.9%; Pred. No. 3.8e-284;
Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Db 289 |||||
GGTCTTCTTCTTAATGAGCCTCTAGATCTATTGTCTGAGAAAGCATTTAAAGCAATT 348
QY 1119 |||||
CCTTCACATCCGAGATCGAATTAACCAAGATGTGGCTTCTGTGCTGCTATGAAGAGGCT 1178
Db 349 |||||
CCTTCACATCCGAGATCGAATTAACCAAGATGTGGCTTCTGTGCTGCTATGAAGAGGCT 408
QY 1179 |||||
CCTGAGATCCTCAGTGGCTCCAAACAATCCCTGACCTCATCGATACAAAATCCCTG 1238
Db 409 |||||
CCTGAGATCCTCAGTGGCTCCAAACAATCCCTGACCTCATCGATACAAAATCCCTG 468
QY 1239 |||||
CAGATCCCGCTCAGTATTTGACCTTTGGCTAATGAATACTTCATGACAAAGCACTCC 1298
Db 469 |||||
CAGATCCCGCTCAGTATTTGACCTTTGGCTAATGAATACTTCATGACAAAGCACTCC 528
QY 1299 |||||
CTGACTGAAATCCGAGACAGTCTTCTGAGACTTGTGAGATGTGTCTTTGTGGTCCCT 1358
Db 529 |||||
CTGACTGAAATCCGAGACAGTCTTCTGAGACTTGTGAGATGTGTCTTTGTGGTCCCT 588
QY 1359 |||||
GCACTGATCAACAGCTCGATATCAACAGATGTGTGACCTGTCTACTTCTATGAGTTT 1418
Db 589 |||||
GCACTGATCAACAGCTCGATATCAACAGATGTGTGACCTGTCTACTTCTATGAGTTT 648
QY 1419 |||||
CGGACCGGCTCAGTGTGTTGAAGACAGAAAGCCGCTTTGTCAAAGCCGACAGCT 1478
Db 649 |||||
CGGACCGGCTCAGTGTGTTGAAGACAGAAAGCCGCTTTGTCAAAGCCGACAGCT 708
QY 1479 |||||
GATGAAGTCCGCTTGTGTGCTGCTGCTTCTGAAAGGGGACATGTATGTTGAA 1538
Db 709 |||||
GATGAAGTCCGCTTGTGTGCTGCTGCTTCTGAAAGGGGACATGTATGTTGAA 768
QY 1539 |||||
GAGGCCAGGAGAGGAGAAATTACTGAGCCGGAAGATGATAAATCTGGGCTACTTT 1598
Db 769 |||||
GAGGCCAGGAGAGGAGAAATTACTGAGCCGGAAGATGATAAATCTGGGCTACTTT 828
QY 1599 |||||
GCTCGAACCGGGAATCTAATGGGAACGACCTGTCTGTGTGGCCAGCTTATATCTGACT 1658
Db 829 |||||
GCTCGAACCGGGAATCTAATGGGAACGACCTGTCTGTGTGGCCAGCTTATATCTGACT 888
QY 1659 |||||
GAGCAGTACCTTCAGCTGACTTGAACATGAGCTCGACAGAGACTCAAGAAACCGCGG 1718
Db 889 |||||
GAGCAGTACCTTCAGCTGACTTGAACATGAGCTCGACAGAGACTCAAGAAACCGCGG 948
QY 1719 |||||
GTGATTTTGTGACAGACACCATCCCTGATCTGTCTGCTCCGACATGCTCCACAGT 1778
Db 949 |||||
GTGATTTTGTGACAGACACCATCCCTGATCTGTCTGCTCCGACATGCTCCACAGT 1008
QY 1779 |||||
CCTCTTCTCTTAATCTTCTCTCTCTCTCCCTCGAGCCCTTTCTTTTGTGCTCCT 1838
Db 1009 |||||
CCTCTTCTCTTAATCTTCTCTCTCTCTCCCTCGAGCCCTTTCTTTTGTGCTCCT 1068
QY 1839 TGA 1841
|||
Db 1069 TGA 1071

Search completed: June 15, 2005, 08:31:05
Job time : 1399.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 9369.39 Seconds
(without alignments)
11160.423 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 2158
Sequence: 1 ccacgcgtccgaaacacag.....aaaaaaaaaaaaaaaa 2158

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_scs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	100.0	2158	6	AR438764 Sequence
2	2158	100.0	2158	6	AX600124 Sequence
3	1953	90.5	2232	6	AX477711 Sequence
4	1917	88.8	2092	6	AX746724 Sequence
5	1917	88.8	2092	9	AK090997 Homo sapi
6	1855	86.0	2047	9	BC069501 Homo sapi
7	1824.6	84.6	2298	9	BC039073 Homo sapi
8	1746	80.9	1746	6	AR438765 Sequence
9	1746	80.9	1746	6	AX600126 Sequence
10	1579.4	73.2	1962	6	AX714052 Sequence
11	1579.4	73.2	1962	9	AK056109 Homo sapi
12	1543.4	71.5	1896	9	BC069548 Homo sapi
13	1425	66.0	1716	6	CO733044 Sequence
14	1386.2	64.2	1857	6	AX583932 Sequence
15	1198.4	55.5	1728	4	AB186392 Canis fam
16	1177.8	53.6	2139	4	AB045377 Felis cat
17	1151.4	53.4	1739	9	AB060873 Macaca fa
18	1041.2	48.2	1728	10	AB186393 Mus muscu
19	634.6	29.4	921	6	AX380871 Sequence

20	530.8	24.6	134482	9	AC007335 Homo sapi
21	530.8	24.6	161419	2	AC136927 Homo sapi
22	530.8	24.6	169792	2	AC147362 Homo sapi
23	526	24.4	173595	9	AC020766 Homo sapi
24	519.2	24.1	90106	2	AC007499 Homo sapi
25	413.8	19.2	2228	10	D50580 Rattus norv
26	408.2	18.9	3391	9	AK095522 Homo sapi
27	406.6	18.8	1846	9	D50579 Homo sapi
28	406.6	18.8	2117	6	CO723952 Sequence
29	406.6	18.8	2191	6	AX330319 Sequence
30	406.6	18.8	2191	6	AX330320 Sequence
31	406.6	18.8	2191	6	AX331596 Sequence
32	406.6	18.8	2191	6	AX411210 Sequence
33	406.6	18.8	2191	6	HSICE Y09616 H.sapiens m
34	406.6	18.8	3155	9	BC032095 Homo sapi
35	405	18.8	1748	9	U60553 Homo sapi
36	401.8	18.6	1831	10	BC028907 Mus muscu
37	401.8	18.6	2139	9	BC053670 Homo sapi
38	400.2	18.5	2114	9	AK000105 Homo sapi
39	400.2	18.5	3824	6	AR528722 Sequence
40	400.2	18.5	3824	6	AX358956 Sequence
41	400.2	18.5	3824	6	AX362449 Sequence
42	400.2	18.5	3824	6	AX454680 Sequence
43	400.2	18.5	3824	6	AX464408 Sequence
44	400.2	18.5	3824	6	AX491158 Sequence
45	400.2	18.5	3824	9	AY358609 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AR438764 AR438764 2158 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 664091.
ACCESSION AR438764 GI:4263739
VERSION AR438764.1 GI:4263739
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2158)
AUTHORS Curtis,R.A.J. and Sillos-Santiago,I.
TITLE 53010, a human carboxylesterase family member and uses thereof
JOURNAL Patent: US 664091-A 1 16-DEC-2003;
FEATURES
Location/Qualifiers
1..2158
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2158; DB 6; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCACGCGTCCGAAACAGCGCTGGAGCAATGAGTAAGCATGTAATATAACGA	60
DB	1	CCACGCGTCCGAAACAGCGCTGGAGCAATGAGTAAGCATGTAATATAACGA	60
QY	61	TTTAAATTTATTTCTTAATAAAGTACGATCCCAATCCACAGGACTTACTTCATCG	120
DB	61	TTTAAATTTATTTCTTAATAAAGTACGATCCCAATCCACAGGACTTACTTCATCG	120
QY	121	CTTCAATGCTGCTTTTCTGATTCACAGCCCTGTTGGACACAGCAGTTCAGGGCA	180
DB	121	CTTCAATGCTGCTTTTCTGATTCACAGCCCTGTTGGACACAGCAGTTCAGGGCA	180
QY	181	AACTGGGCTTCTGCTGAAGGCGACAGAGAACACAGGCTGGAGTTCAGGGCA	240
DB	181	AACTGGGCTTCTGCTGAAGGCGACAGAGAACACAGGCTGGAGTTCAGGGCA	240
QY	241	AGCAAGTCACTGTGCTGGAGGCGCTGCTGTGAGCGTTCCTCGAGTCCCTTTG	300
DB	241	AGCAAGTCACTGTGCTGGAGGCGCTGCTGTGAGCGTTCCTCGAGTCCCTTTG	300

Db 241 AGCAAGTCATGTGTGGGAGGCCCTGAGCCTGTGAAAGTGTCTCTGGAGTCCCTTTG 300
 Qy 301 CTGCTCCCGCGCTGGAGATCCCTGGATTATAGAAACCGGACCTTGCATTCGCTGGAGATA 360
 Db 301 CTGCTCCCGCGCTGGAGATCCCTGGATTATAGAAACCGGACCTTGCATTCGCTGGAGATA 360
 Qy 361 ACTTGGAGAGACCACTCTACCTTAATTTGTGCTCCAGAACTGAGAGTGTGCTCT 420
 Db 361 ACTTGGAGAGACCACTCTACCTTAATTTGTGCTCCAGAACTGAGAGTGTGCTCT 420
 Qy 421 TAGATCAACACATCTCAAGTGTGATTAACCGAAATTCGAGGTGCAGAAAGTGTCTCT 480
 Db 421 TAGATCAACACATCTCAAGTGTGATTAACCGAAATTCGAGGTGTGAGAGTGTCTCT 480
 Qy 481 ACCGAAATCTATAGGCTGTGCCAGAGCGATTAACGGCTCAAGCTCCCGTCTTGATGT 540
 Db 481 ACCGAAATCTATAGGCTGTGCCAGAGCGATTAACGGCTCAAGCTCCCGTCTTGATGT 540
 Qy 541 GGTTCACGAGAGGTGCTTCAAGACTGTGCTCAAGCTTCATCTTGTGATGGGTCCGCTGTG 600
 Db 541 GGTTCACGAGAGGTGCTTCAAGACTGTGCTCAAGCTTCATCTTGTGATGGGTCCGCTGTG 600
 Qy 601 CTGCTCAATGAGACGT 660
 Db 601 CTGCTCAATGAGACGT 660
 Qy 661 TCAACAATGAGATCAGATGTCTCCGAGGAGCTGGGCTTCAAGGACGAGGTGTGTGTGT 720
 Db 661 TCAACAATGAGATCAGATGTCTCCGAGGAGCTGGGCTTCAAGGACGAGGTGTGTGTGT 720
 Qy 721 TGTCTGT 780
 Db 721 TGTCTGT 780
 Qy 781 TTGGGAGTCCGAGGAGGCTTAAGTGTCTTAAGTGTCTTAAGTGTCTTAAGTGTCTTAAG 840
 Db 781 TTGGGAGTCCGAGGAGGCTTAAGTGTCTTAAGTGTCTTAAGTGTCTTAAGTGTCTTAAG 840
 Qy 841 GCTTATTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTCACTGAGAGGCC 900
 Db 841 GCTTATTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTCACTGAGAGGCC 900
 Qy 901 ATGATTAAGAGAGTGAAGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Db 901 ATGATTAAGAGAGTGAAGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Qy 961 CAGACTGTGAGGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 Db 961 CAGACTGTGAGGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 Qy 1021 TCAGGCAAAAACAAAGTCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 Db 1021 TCAGGCAAAAACAAAGTCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 Qy 1081 CTCTAGATCTATGTGTGAGAAAGCATTAAGCAATTCCTTCACTGAGTGTGTGTGTGTGT 1140
 Db 1081 CTCTAGATCTATGTGTGAGAAAGCATTAAGCAATTCCTTCACTGAGTGTGTGTGTGTGT 1140
 Qy 1141 ACCAAGAGT 1200
 Db 1141 ACCAAGAGT 1200
 Qy 1201 ACAAGTCCCTTGCCTGCATCTGATTAACAAACATCTGACATCCGCTCAGATTTTGC 1260
 Db 1201 ACAAGTCCCTTGCCTGCATCTGATTAACAAACATCTGACATCCGCTCAGATTTTGC 1260
 Qy 1261 ACCTTGT 1320
 Db 1261 ACCTTGT 1320
 Qy 1321 TTCTGGAATGCTGTGAGATGT 1380
 Db 1321 TTCTGGAATGCTGTGAGATGT 1380

Qy 1381 ACAGAGATCTGTGTGACCTGTCTATCTTCTATGAGTTTCGGACCGGCTCAGTGTGTG 1440
 Db 1381 ACAGAGATCTGTGTGACCTGTCTATCTTCTATGAGTTTCGGACCGGCTCAGTGTGTG 1440
 Qy 1441 AAGACAGAGAGCGGCTTTTGTCAAAAGCGACCAAGCTGATGAATCCGCTTGTGTGTG 1500
 Db 1441 AAGACAGAGAGCGGCTTTTGTCAAAAGCGACCAAGCTGATGAATCCGCTTGTGTGTG 1500
 Qy 1501 GTGTGCTCTTCTGAAAGGGGAGCATTTGTATGTTTGAAGAGCCACGAGAGAGAGAA 1560
 Db 1501 GTGTGCTCTTCTGAAAGGGGAGCATTTGTATGTTTGAAGAGCCACGAGAGAGAGAA 1560
 Qy 1561 TACTAGCCGAGAGATGATGAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Db 1561 TACTAGCCGAGAGATGATGAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
 Qy 1621 GGAAGACCTGTCTTGT 1680
 Db 1621 GGAAGACCTGTCTTGT 1680
 Qy 1681 TGAACATGAGCTGTGACAGAGACTCAAAAGACCGGGGTGATTTTGAACAGAGCA 1740
 Db 1681 TGAACATGAGCTGTGACAGAGACTCAAAAGACCGGGGTGATTTTGAACAGAGCA 1740
 Qy 1741 TCCCTGT 1800
 Db 1741 TCCCTGT 1800
 Qy 1801 TCTCTCTCTCAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 Db 1801 TCTCTCTCTCAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 Qy 1861 TTGAGTTCCCTTCTCTCCATTAATTCCTCCGCAATCATTAAGTGTGTGTGTGTGTGT 1920
 Db 1861 TTGAGTTCCCTTCTCTCCATTAATTCCTCCGCAATCATTAAGTGTGTGTGTGTGTGT 1920
 Qy 1921 GCTGCTTCTTAAGGAGTCTTGTCAAAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
 Db 1921 GCTGCTTCTTAAGGAGTCTTGTCAAAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
 Qy 1981 GGAATGATCTTCAAGATCTTCTTCAACATCAAAAGTGTGTGTGTGTGTGTGTGTGTGT 2040
 Db 1981 GGAATGATCTTCAAGATCTTCTTCAACATCAAAAGTGTGTGTGTGTGTGTGTGTGTGT 2040
 Qy 2041 CAAGATTTCTTCAATTAATTTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
 Db 2041 CAAGATTTCTTCAATTAATTTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
 Qy 2101 TGTACTCATTAAGATTAATTAAGATTAATTAAGATTAATTAAGATTAATTAAGATTA 2158
 Db 2101 TGTACTCATTAAGATTAATTAAGATTAATTAAGATTAATTAAGATTAATTAAGATTA 2158
 RESULT 2
 AX600124 2158 bp DNA linear PAT 14-FEB-2003
 LOCUS
 DEFINITION Sequence 1 from Patent WO0250256.
 AX600124
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Curtis, R.A. and Silos-Santiago, I.
 53010, a novel human carboxylesterase family member and uses
 thereof
 Patent: WO 0250256-A 1 27-JUN-2002;
 Millennium Pharmaceuticals, Inc. (US)
 JOURNAL
 FEATURES
 source 1..2158

CDS

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 96..1841
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD66643.1"
 /db_xref="GI:28400191"
 /translation="MPDGLTSSASQWCFLLIOLPLIGHROMKGTSPASRGPORTRLG
 WIOKQVYVLSGPPVNVNPLGVPAAPPLGSRFTNPPQASPMDLRBRATSYDNCLQ
 NSEMLLDQHLKVTYPRKGVSECLYLTAPAHADTSKLPVLPFGAFKTSQ
 SIFDSSALAAVEDVAVVWVQYRLGIFGETTWDOHAPGWAKDOVAALSVQKILF
 FGDPSSVYILFEGSAGIISVSSLILSPMAKGLFKHAIMSGVAILPYLEADHYKESD
 LQVAFPCGNMNSDEBALRCRTKPSKELTISOKTSGFTVVDGAFPPNPDLDS
 OKAFKAIPIIIGVNNHGGFLLPMKEAPRLLSGSNKSLALHILIONIHLIPQYLHLVA
 NEYFHKSLTETIRDSLIDLSDVFPVVPALITARYHNDAGPYFYERHRPQCPED
 TKPAPVADHADVPRFEGATLKQDITVFEGBATEERKILSKMKMTATYFRTGSPN
 GNDLSLWPAYNLTQYLOLDLNLMSLQRLKEPRVDFWSTIPLILSASDMLSPSL
 TFLSLQPFPPFCAP"

ORIGIN

Query Match 100.0%; Score 2158; DB 6; Length 2158;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CCAGCGTCCGAAAAAGAGCCCTGGAGAGCATGTGGAATGATGATATTAAGAAAGAA 60
DB 1 CCAGCGTCCGAAAAAGAGCCCTGGAGAGCATGTGGAATGATGATATTAAGAAAGAA 60
OY 61 TTTAAAAATTATCTTAATAAAGTACGAATCCCATGCGACAGAGGACTTACTCATCTG 120
DB 61 TTTAAAAATTATCTTAATAAAGTACGAATCCCATGCGACAGAGGACTTACTCATCTG 120
OY 121 CTTCACATGTCCTTTTCTCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGAA 180
DB 121 CTTCACATGTCCTTTTCTCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGAA 180
OY 181 AAACGGGGCTTCTGTGAAGGGCCACAGAGAACAGAGGCTGGAGTGGATTCAGGGCA 240
DB 181 AAACGGGGCTTCTGTGAAGGGCCACAGAGAACAGAGGCTGGAGTGGATTCAGGGCA 240
OY 241 AGCAAGTCACTGTGCTGGAGAACCCCTGTGCTGTGAACGTGTCTCGAGATCCCTTTG 300
DB 241 AGCAAGTCACTGTGCTGGAGAACCCCTGTGCTGTGAACGTGTCTCGAGATCCCTTTG 300
OY 301 CTGCTCCCCCGCTGGGATCCCTGCGATTACGAACCCGAGCCTGCATGCCCTGGATA 360
DB 301 CTGCTCCCCCGCTGGGATCCCTGCGATTACGAACCCGAGCCTGCATGCCCTGGATA 360
OY 361 ACTTGCGAGAACCCCTCTACCCCTAATTGTGCTCCAGAACCTCAGATGGGCTGCT 420
DB 361 ACTTGCGAGAACCCCTCTACCCCTAATTGTGCTCCAGAACCTCAGATGGGCTGCT 420
OY 421 TAGATCAACACATGCTCAAGTGTATTACCCGAATTCCGAGTGTCAAGAGCTGCTCT 480
DB 421 TAGATCAACACATGCTCAAGTGTATTACCCGAATTCCGAGTGTCAAGAGCTGCTCT 480
OY 481 ACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCCAGCTCCCGCTTGGTGT 540
DB 481 ACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCCAGCTCCCGCTTGGTGT 540
OY 541 GGTTCGCCGAGAGTCTTCAAGACTGGCTCAGCCCTCCATCTTTGATGGGCTCGCCCTG 600
DB 541 GGTTCGCCGAGAGTCTTCAAGACTGGCTCAGCCCTCCATCTTTGATGGGCTCGCCCTG 600
OY 601 CTGCTATAGAGACGTGCTGTGTGTGTGCTCAGTACCGGCTAGAGATATTTGGTTCT 660
DB 601 CTGCTATAGAGACGTGCTGTGTGTGTGCTCAGTACCGGCTAGAGATATTTGGTTCT 660
OY 661 TCACCAATGGATCAGATGCTCCGGGAACTGGGCTTCAAGACACAGTGGCTGCTC 720
DB 661 TCACCAATGGATCAGATGCTCCGGGAACTGGGCTTCAAGACACAGTGGCTGCTC 720

```

```

OY 721 TGTCTGGGTCCAGAGAACATCGAGTTCTTGGTGGGACCCAGCTCTGTGACCATCT 780
DB 721 TGTCTGGGTCCAGAGAACATCGAGTTCTTGGTGGGACCCAGCTCTGTGACCATCT 780
OY 781 TTGGGAGTCCCGGGAGGCGATTAAGTGTTCATGCTTATCTGTCTCCCATGGCCAAAG 840
DB 781 TTGGGAGTCCCGGGAGGCGATTAAGTGTTCATGCTTATCTGTCTCCCATGGCCAAAG 840
OY 841 GCTTATTTCCAAAGCCATCATGAGAGTGGGGGCCATCCCTTACCTGAGAGGCC 900
DB 841 GCTTATTTCCAAAGCCATCATGAGAGTGGGGGCCATCCCTTACCTGAGAGGCC 900
OY 901 ATGATTATGAGAGAGTGAAGACCTGACAGTGGTTGCACTTCTGTGTGAACAATGCGT 960
DB 901 ATGATTATGAGAGAGTGAAGACCTGACAGTGGTTGCACTTCTGTGTGAACAATGCGT 960
OY 961 CAGACTCTGAGGCCCTGCTGAGAGTGCCTGAGAGCAAAACCTTCCAAAGAGCTGACCC 1020
DB 961 CAGACTCTGAGGCCCTGCTGAGAGTGCCTGAGAGCAAAACCTTCCAAAGAGCTGACCC 1020
OY 1021 TCAGCCAGAAACAAAGTCTTCACTCGAGTGGTTGATGGTCTTCTTCCAAATGAGC 1080
DB 1021 TCAGCCAGAAACAAAGTCTTCACTCGAGTGGTTGATGGTCTTCTTCCAAATGAGC 1080
OY 1081 CTCTAGATCTATTTGCTCAGAAAGCATTTAAAGCAATTCCTTCCATCATCGAGTCAATA 1140
DB 1081 CTCTAGATCTATTTGCTCAGAAAGCATTTAAAGCAATTCCTTCCATCATCGAGTCAATA 1140
OY 1141 ACCAGAGTGTGGCTTCTGTGCTGCTGCTGATGAGAGGCTCTGAGATCTCAAGTGGCTCA 1200
DB 1141 ACCAGAGTGTGGCTTCTGTGCTGCTGCTGATGAGAGGCTCTGAGATCTCAAGTGGCTCA 1200
OY 1201 ACAAATCCCTTGGCCCTCCATCGATTAACAAACATCCTGACATCCCGGCTCAGTATTTGC 1260
DB 1201 ACAAATCCCTTGGCCCTCCATCGATTAACAAACATCCTGACATCCCGGCTCAGTATTTGC 1260
OY 1261 ACCTGTGGCTAATGAAATCTTCCATGACAAAGCACTCCCTGACGTAATCCGAGACGTC 1320
DB 1261 ACCTGTGGCTAATGAAATCTTCCATGACAAAGCACTCCCTGACGTAATCCGAGACGTC 1320
OY 1321 TTCTGAGCTTGTGAGAGTGTCTTTGTGTGCTCTGCACTGATCAGACTCGATATTC 1380
DB 1321 TTCTGAGCTTGTGAGAGTGTCTTTGTGTGCTCTGCACTGATCAGACTCGATATTC 1380
OY 1381 AAGAGATGCTGTGACACTGTGCTACTTGAAGTTTCGAGACGGGCTCAGTGTGTTG 1440
DB 1381 AAGAGATGCTGTGACACTGTGCTACTTGAAGTTTCGAGACGGGCTCAGTGTGTTG 1440
OY 1441 AAGACAGAAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAGTCCGTTTGTGTTG 1500
DB 1441 AAGACAGAAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAGTCCGTTTGTGTTG 1500
OY 1501 GTGTGCTTCTTGAAGGGGACATTTGTTATGTTGGAAGAGCCACGAGAGAGAGAGT 1560
DB 1501 GTGTGCTTCTTGAAGGGGACATTTGTTATGTTGGAAGAGCCACGAGAGAGAGAGT 1560
OY 1561 TACTGAGCCGGAAGATGATGAATACTGGGCTACCTTTGCTGGAACCGGGAATCCAAAG 1620
DB 1561 TACTGAGCCGGAAGATGATGAATACTGGGCTACCTTTGCTGGAACCGGGAATCCAAAG 1620
OY 1621 GGAACGACCTGTCTGTGAGCAGCTTAATCTGACAGAGTACCTCCAGCTGAGCT 1680
DB 1621 GGAACGACCTGTCTGTGAGCAGCTTAATCTGACAGAGTACCTCCAGCTGAGCT 1680
OY 1681 TGAACATGAGCTTGGAGACAGAGCTCAAAAGAACCCGGGTGATTTTGGACACAGACA 1740
DB 1681 TGAACATGAGCTTGGAGACAGAGCTCAAAAGAACCCGGGTGATTTTGGACACAGACA 1740
OY 1741 TCCCTCTGATCTGTGCTGCTCCGACATGCTCACAAGTCTCTTCTTCCCTTAACCTTCC 1800
DB 1741 TCCCTCTGATCTGTGCTGCTCCGACATGCTCACAAGTCTCTTCTTCTTAACTTCTCC 1800
OY 1801 TCTCTCTCTCAGCTTCTTCTTCTTGTGCTCTTGAAGATTAATCTTCTGCTGAT 1860

```


Db 1801 TCTCTCTCTCCAGCCTTTCTTTCTTTTGTGGCTCTGGAAAGTATATCTTCTGTGAT 1860

Qy 1861 TTGTGTTTCCCTTCTCTCTCCAAATATTTCTCCGGAATCATAGCTTCTTTCTGAGCTCA 1920

Db 1861 TTGTGTTTCCCTTCTCTCTCCAAATATTTCTCCGGAATCATAGCTTCTTTCTGAGCTCA 1920

Qy 1921 GCTGCTTTCTAAGGGAGTCCCTGGAAAACAAGCTGCTTTCCGTGATATTTATAGACTTA 1980

Db 1921 GCTGCTTTCTAAGGGAGTCCCTGGAAAACAAGCTGCTTTCCGTGATATTTATAGACTTA 1980

Qy 1981 GGAATGATCTTACAGAAATCTTTTCAACATCTAAAAGTGCATTTGTCTGGAAGCAA 2040

Db 1981 GGAATGATCTTACAGAAATCTTTTCAACATCTAAAAGTGCATTTGTCTGGAAGCAA 2040

Qy 2041 CAAGATTTCTTCAATAAATTTTGGAAAGAGGCTGGCTTATTTAGTTGTCAATATAATGTGTT 2100

Db 2041 CAAGATTTCTTCAATAAATTTTGGAAAGAGGCTGGCTTATTTAGTTGTCAATATAATGTGTT 2100

Qy 2101 TGTAACTCATATGAAAATAAATCAGAAATGTAAATAGGAAAAAATTTTTTTTTTTTTTTT 2158

Db 2101 TGTAACTCATATGAAAATAAATCAGAAATGTAAATAGGAAAAAATTTTTTTTTTTTTTTT 2158

RESULT 3					
AX477711					
LOCUS	AX477711	2232 bp	DNA	linear	PAT 12-AUG-2002
DEFINITION	Sequence 23 from Patent WO0246426.				
ACCESSION	AX477711				
VERSION	AX477711.1	GI:22216866			
KEYWORDS	Homo sapiens (human)				
SOURCE					

ORIGIN

Query Match	90.5%	Score 1953;	DB 6;	Length 2232;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1959; Conservative	0;	Mismatches	10;	Indels 0;

QY	176	GGGAAAAA	CTGGCCCT	TTCTGCTG	AGGGCC	CACAGAA	CACCAAGCTGGATG	ATTTCA	235	
Db	264	GCTCAGAAG	GGGCTTT	CTGCTGA	AGGGCC	CACAGAA	CACCAAGCTGGATG	ATTTCA	323	
QY	236	GGGCAAGCA	AGTCACTGTG	CTGGGA	AGCCCTGTG	CGCTTGAA	CGATTTCTT	CGAGTCCC	295	
Db	324	GGGCAAGCA	AGTCACTGTG	CTGGGA	AGCCCTGTG	CGCTTGAA	CGATTTCTT	CGAGTCCC	383	
QY	296	CTTGTGCT	CCCCCG	CTGGAGT	CCCTCG	ATTTACGA	ACCGGCGCTGAT	CGCCCTG	355	
Db	384	CTTGTGCT	CCCCCG	CTGGAGT	CCCTCG	ATTTACGA	ACCGGCGCTGAT	CGCCCTG	443	
QY	356	GGATTA	ACTTGGAG	AAAGCC	ACCTCTTA	CCCTAA	TTTGTG	CGCTCCGAA	CTCAGATG	415
Db	444	GGATTA	ACTTGGAG	AAAGCC	ACCTCTTA	CCCTAA	TTTGTG	CGCTCCGAA	CTCAGATG	503
QY	416	GCTCTG	ATGAACA	CAGTCT	CAGAGT	GCATTA	CCCGAA	ATTGGAGT	GTCAAGAA	475

Dp	504	GCTCTTAATCAACAACATGCTCAAGGTGCAATTACCCGAAATTTGCGAGTGTCAAGAACTG	563
Qy	476	CCTTCACTGTAACATCTATATGCGCTGCGCCAGCGGATACAGGCTCCAAAGCTCCCGCTTT	533
Dp	564	CCTTCACTGTAACATCTATATGCGCTGCGCCAGCGGATACAGGCTCCAAAGCTCCCGCTTT	622
Qy	536	GATGATGATCCCAAGAGGTGCTTCAAGACTGCGTCAAGCCTTCCATCTTTGATGGGTCCG	599
Dp	624	GGTGTGGTTCCCAAGAGGTGCTTCAAGACTGCGTCAAGCCTTCCATCTTTGATGGGTCCG	683
Qy	596	CTTGAGCTGCTATAGAGACGTGCTGTTGTGTGTCAGTATACCGGCTAGGAATATTGG	655
Dp	684	CTTGAGCTGCTATAGAGACGTGCTGTTGTGTGTCAGTATACCGGCTAGGAATATTGG	743
Qy	656	TTTTTTTCAACAATNGGATCAAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGATGC	715
Dp	744	TTTTTTTCAACAATNGGATCAAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGATGC	803
Qy	716	TGCTCTGCTCTGGGTCAGAGAAACATCGAGTTCTTCGATGGGAGACCCGAGCTCTGAGC	775
Dp	804	TGCTCTGCTCTGGGTCAGAGAAACATCGAGTTCTTCGATGGGAGACCCGAGCTCTGAGC	863
Qy	776	CATCTTTGGGGAATCCGCGGAGCCATAAGTGTTCATGCTTATACTGTCCCACTGAC	835
Dp	864	CATCTTTGGGGAATCCGCGGAGCCATAAGTGTTCATGCTTATACTGTCCCACTGAC	923
Qy	836	CAAAAGCTTATTTCCACAAAGCCATCATGTGAGATGGGGGTGGCCATATCCCTTACTGGA	895
Dp	924	CAAAAGCTTATTTCCACAAAGCCATCATGTGAGATGGGGGTGGCCATATCCCTTACTGGA	983
Qy	896	GGGCCATGATATAGAGAAGTGAAGACCTGCAAGTGTGTTCACATTTCTGTGTAAACA	955
Dp	984	GGGCCATGATATAGAGAAGTGAAGACCTGCAAGTGTGTTCACATTTCTGTGTAAACA	1041
Qy	956	TGCGTCAGACTCTGAGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGACTGCT	1011
Dp	1044	TGCGTCAGACTCTGAGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGACTGCT	1103
Qy	1016	GACCTCAGCCCAAGAAAACAAGTCTTTCATCTCGAGTGTGTGATGGTCTTTCTCTTA	1072
Dp	1104	GACCTCAGCCCAAGAAAACAAGTCTTTCATCTCGAGTGTGTGATGGTCTTTCTCTTA	1162
Qy	1076	TGAGGCTCTAGATCTATGTGCTCAAGAAAGCATTTAAAGCAATCTTCCATCATCTGAGT	1133
Dp	1164	TGAGGCTCTAGATCTATGTGCTCAAGAAAGCATTTAAAGCAATCTTCCATCATCTGAGT	1223
Qy	1136	CAATTAACAAGATGTGCTTCTGCTGCTCATGAAAGAGGCTCTGAGATCTCTAGAG	1199
Dp	1224	CAATTAACAAGATGTGCTTCTGCTGCTCATGAAAGAGGCTCTGAGATCTCTAGAG	1283
Qy	1196	CTCCAAACAAGTCCCTTGCGCTTCATCTGATACAAAACATCTGTGCACATCCGCGCTCA	1255
Dp	1284	CTCCAAACAAGTCCCTTGCGCTTCATCTGATACAAAACATCTGTGCACATCCGCGCTCA	1343
Qy	1256	TTTGACCTTGCTATGTAATTAATCTTCATGACAAAGACCTCCGATGTAATCCGAGA	1315
Dp	1344	TTTGACCTTGCTATGTAATTAATCTTCATGACAAAGACCTCCGATGTAATCCGAGA	1403
Qy	1316	CAGTCTTCTGCACTTGCTTGGAATGTGTCTTTGTGTGTCTCTGCACTGATATCAGCTGC	1375
Dp	1404	CAGTCTTCTGCACTTGCTTGGAATGTGTCTTTGTGTGTCTCTGCACTGATATCAGCTGC	1463
Qy	1376	ATTATCAAGAGATCTGCTGTGCACTGTCTTCTATAGAGTTTCCGACCCGAGCTCAGTG	1435
Dp	1464	ATTATCAAGAGATCTGCTGTGCACTGTCTTCTATAGAGTTTCCGACCCGAGCTCAGTG	1523
Qy	1436	CTTTGAAGACACGAGCCGCTTTTGTCAAAAGCCGACACAGCTGATGATGCTCCCTTGT	1495
Dp	1524	CTTTGAAGACACGAGCCGCTTTTGTCAAAAGCCGACACAGCTGATGATGCTCCCTTGT	1583
Qy	1496	GTTCGGGTGGTCTCTTCTGAAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGA	1555
Dp	1584	GTTCGGGTGGTCTCTTCTGAAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGA	1643

Db 1218 TTGCACTGTGGCTAATGATATATCTTCACAGACAGACCTCCGACTGAAATCCAGAA 1277
 Oy 1316 CAGCTTCTGACCTGCTGAGATGATGTTCTTTGGTCCCTGACATGATCACAGCTCG 1375
 Db 1278 CAGCTTCTGACCTGCTGAGATGATGTTCTTTGGTCCCTGACATGATCACAGCTCG 1337
 Oy 1376 ATATACAGAGATGCTGAGACCTGCTGATCTTCAATGATGTTGGACCGGCTCTGAG 1435
 Db 1338 ATATACAGAGATGCTGAGACCTGCTGATCTTCAATGATGTTGGACCGGCTCTGAG 1397
 Oy 1436 CTTTGAAGACGGAAGCCGCTTTTGTCAAAAGCCGACCGCTGATGAGTCCGCTTGT 1495
 Db 1398 CTTTGAAGACGGAAGCCGCTTTTGTCAAAAGCCGACCGCTGATGAGTCCGCTTGT 1457
 Oy 1496 GTTGGTGGTGGCTTCCGAAAGGGGGGACATTTGATATTTGAAAGAGCCACGAGAGGA 1555
 Db 1458 GTTGGTGGTGGCTTCCGAAAGGGGGGACATTTGATATTTGAAAGAGCCACGAGAGGA 1517
 Oy 1556 GAAGTTACTAGCCGGAAGATGATGAAATATCTGGGCTACCTTGTCTGAAACCGGAATCC 1615
 Db 1518 GAAGTTACTAGCCGGAAGATGATGAAATATCTGGGCTACCTTGTCTGAAACCGGAATCC 1577
 Oy 1616 TAAATGGGAAGACCTGTCTGTGGCCAGCTTATCTGATGAGACGATCCTCCAGCT 1675
 Db 1578 TAAATGGGAAGACCTGTCTGTGGCCAGCTTATCTGATGAGACGATCCTCCAGCT 1637
 Oy 1676 GAACTTGAACATGAGCTCGGACAGAGACTCAAAAGAACCGGGGGGATTTTGGACCG 1735
 Db 1638 GAACTTGAACATGAGCTCGGACAGAGACTCAAAAGAACCGGGGGGATTTTGGACCG 1697
 Oy 1736 CAGCAATCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTCTTCTCTCTTAAC 1795
 Db 1698 CAGCAATCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTCTTCTCTCTTAAC 1757
 Oy 1796 TTTCT 1855
 Db 1758 TTTCT 1817
 Oy 1856 GTGATTTGGTTCCT 1915
 Db 1818 GTGATTTGGTTCCT 1877
 Oy 1916 GCTCAGCTGCT 1975
 Db 1878 GCTCAGCTGCT 1937
 Oy 1976 ACTTAGAATGATCT 2035
 Db 1938 ACTTAGAATGATCT 1997
 Oy 2036 GGCAACAGATTTCT 2095
 Db 1998 GGCAACAGATTTCT 2057
 Oy 2096 GGTTTTGTACTCATGATGAAATTAATCAGATGT 2130
 Db 2058 GGTTTTGTACTCATGATGAAATTAATCAGATGT 2092

RESULT 5
 AK090997
 LOCUS 2092 bp mRNA linear PRI 30-JAN-2004
 DEFINITION Homo sapiens cDNA FLJ33678 fis' clone BRAHM2002191, weakly similar
 to FATTY ACYL-CoA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
 AK090997
 ACCESSION AK090997.1 GI:21749268
 VERSION oligo capping; fis (full insert sequence).
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1

AUTHORS

Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Nakita, H.,
 Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
 Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagaetsuma, M.,
 Shiratori, A., Sudo, H., Hosokawa, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Fujiyama, T.,
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
 Tanikawa, M., Yamazaki, M., Nishimura, K., Ishihara, T., Yamashita, H.,
 Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
 Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
 Yoshida, M., Houta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,
 Hara, H., Tanase, T., Nomura, Y., Togiyama, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Muesahino, K., Yuuki, H., Oshima, A.,
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
 Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satch, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
 Kunigai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
 Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
 Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
 Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
 Noguchi, S., Itoh, T., Shigeta, K., Senda, T., Matsunura, K.,
 Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
 Oyama, M., Hara, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
 Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
 Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R.,
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
 Complete sequencing and characterization of 21,243 full-length
 human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 2
 14702039

JOURNAL
 PUBMED
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

FEATURES
 source

CDS

1. 2092
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAMH2002191"
 /tissue_type="Brain"
 /clone_lib="BRAMH2"
 /note="Cloning vector: pME18SFL13"
 394..1803
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC03565.1"
 /db_xref="GI:21749268"
 /translation="MLKHYPRKPVSEDCILYNIYAPAHADTGSKLPLVWVFGAGAK
 TGSASIPDGSALAAVEDLVVVOYRIGLIGFETTWQHAFGNNAFQDVAAALSWQK

ORIGIN

NIEFPGDPSVITFGESAGAIYSVSLILSPMAKGLFKHAKMESVAIIPYLEADYE
 KSEBIOVAFKCNNA SDSBALRLCRITLRSOKRSPFRVVDGAFPEPBL
 DLISOKARAPSIITGVN SHCGPLLPKHEAPELISGSKSLAHLIONIIHIPROYL
 HUVANEYPHDKSLSLEIRDSLLDLGDPVFPVPAIITRHYRDAGAPYTFEPRRPO
 CFEDTKPAFVADHADVRFVFGAFNLGDVMEGFAVEERLELSRKMKWATPART
 GPNPNDLSLWPAVNLTOQYLOLDLNMSLGRLKRPVEFWTSTIPLILSADMHSR
 LSLITFLSLQPPFFFCAP"

Query Match 88.8%; Score 1917; DB 9; Length 2092;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;

186 GGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAAGGCAAGAA 245
 138 GGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAAGGCAAGAA 197
 246 GTCACTGTGTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGGAGTCCCTTTGCTGCT 305
 198 GTCACTGTGTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGGAGTCCCTTTGCTGCT 257
 306 CCCCCGCTGGGATCCCTGCGATTACGAACCGGACGCTGCATCGCCCTGGGATTAATTG 365
 258 CCCCCGCTGGGATCCCTGCGATTACGAACCGGACGCTGCATCGCCCTGGGATTAATTG 317
 366 CGAGAGGACCTCTCTACCTTAATTT-----GTGGCTCCAGAACTCGAGAGTGGCT 415
 318 CGAGAGGACCTCTCTACCTTAATTTGTAAACAGAGGCTCTCGAGAACTCGAGAGTGGCT 377
 416 GCTCTTAAGTCAACATGCTCAAGTGCATTAACCGAAATTCGAGTGTCAAGAGACTG 475
 378 GCTCTTAAGTCAACATGCTCAAGTGCATTAACCGAAATTCGAGTGTCAAGAGACTG 437
 476 CCTCTAAGTCAACATGCTCAAGTGCCTCCAGGCTCAAGGCTCAAGGCTCCCTGCTT 535
 438 CCTCTAAGTCAACATGCTCAAGTGCCTCCAGGCTCAAGGCTCAAGGCTCCCTGCTT 497
 536 GGTGTGTTCCAGAGAGGTGCTTCAAGATGAGTCAAGCTCCATCTTGAAGTGGTCCG 595
 498 GGTGTGTTCCAGAGAGGTGCTTCAAGATGAGTCAAGCTCCATCTTGAAGTGGTCCG 557
 596 CCTGCTGCTCATAGAGAGCTGTGCTGTGTGCTGTCAAGTACCGGCTAGAGAAATTTTG 655
 558 CTTGGCTGCTCATAGAGAGCTGTGCTGTGTGCTGTCAAGTACCGGCTAGAGAAATTTTG 617
 656 TTTCTTCAACATGAGGATCAAGATGCTCCGAGAACTGGGCTTCAAGAACAGGTGGC 715
 618 TTTCTTCAACATGAGGATCAAGATGCTCCGAGAACTGGGCTTCAAGAACAGGTGGC 677
 716 TGTCTGTGCTGGGCTCCAGAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTAC 775
 678 TGTCTGTGCTGGGCTCCAGAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTAC 737
 776 CATCTTTGGAGAGTCCGCGGAGCCATAGTGTCTTCAAGTCTTCAAGTCTCCAGTGC 835
 738 CATCTTTGGAGAGTCCGCGGAGCCATAGTGTCTTCAAGTCTTCAAGTCTCCAGTGC 797
 836 CAAGAGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACCTGGA 895
 798 CAAGAGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACCTGGA 857
 896 GGGCCATGATATATAGAGAGTGAAGTCAAGTGTGTTCAACATCTTGTGTTAAAGAA 955
 858 GGGCCATGATATATAGAGAGTGAAGTGTGTTCAACATCTTGTGTTAAAGAA 917
 956 TGGCTCAAGACTCTGAGGCTCTGCTGAGGCTGCTGAGACAAACCTCTCAAGAGTCTCT 1015
 918 TGGCTCAAGACTCTGAGGCTCTGCTGAGGCTGCTGAGACAAACCTCTCAAGAGTCTCT 977
 1016 GACCTTCAGCCAGAAACAAAGTCTTTCATCTGAGTGTGATGTTGCTTCTTCTTAA 1075
 978 GACCTTCAGCCAGAAACAAAGTCTTTCATCTGAGTGTGATGTTGCTTCTTCTTAA 1037

1076 TGAGCTCTAGATCTATTTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGT 1135
 1038 TGAAGCTCTAGATCTATTTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGT 1097
 1136 CAATTAACAGAGTGTGCTTCTGCTGCTCTATGAGAGAGGCTCTGAGATCTCAGTGG 1195
 1098 CAATTAACAGAGTGTGCTTCTGCTGCTCTATGAGAGAGGCTCTGAGATCTCAGTGG 1157
 1196 CTCCAAAGAGTCCCTTGGCCCTCCATCTGATCAAAAACATCTGCAATCCCGCCCTAGTA 1255
 1158 CTCCAAAGAGTCCCTTGGCCCTCCATCTGATCAAAAACATCTGCAATCCCGCCCTAGTA 1217
 1256 TTTGACCTTGTGCTTAATGAAATCTTCCATGACAGAGCACTCCCTGACTGAATCCGAGA 1315
 1218 TTTGACCTTGTGCTTAATGAAATCTTCCATGACAGAGCACTCCCTGACTGAATCCGAGA 1277
 1316 CAGTCTTGTGAATCTTGTGAGATGTTCTTTGTGTTCTTGTGCTCCCTGCACTGATCAGCTG 1375
 1278 CAGTCTTGTGAATCTTGTGAGATGTTCTTTGTGTTCTTGTGCTCCCTGCACTGATCAGCTG 1337
 1376 AATACAGAGATGCTGTGAGACCTGTCTACTTCTATAGATTGGGCAACGGCCCTAGTG 1435
 1338 AATACAGAGATGCTGTGAGACCTGTCTACTTCTATAGATTGGGCAACGGCCCTAGTG 1397
 1436 CTTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1495
 1398 CTTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1457
 1496 GTTCGGTGTGCTCTTCTGAAGGGGGAATTTGTATGTTTCAAGAGAGCCACGAGAGAGA 1555
 1458 GTTCGGTGTGCTCTTCTGAAGGGGGAATTTGTATGTTTCAAGAGAGCCACGAGAGAGA 1517
 1556 GAAGTATCTGAGCCGGGAAGATGAATTAATGAGGCTACTTGTGCTCAACCGGGAATCC 1615
 1518 GAAGTATCTGAGCCGGGAAGATGAATTAATGAGGCTACTTGTGCTCAACCGGGAATCC 1577
 1616 TAATGGAGACAGCTGTCTCTGTGGCCAGTTATATGATGACTGAGAGTACTTCACT 1675
 1578 TAATGGAGACAGCTGTCTCTGTGGCCAGTTATATGATGACTGAGAGTACTTCACT 1637
 1676 GGAATTGAACATGAGCTCTGAGACAGAGACTCAAAAGAACCGGGGTGATATTTTGGACAG 1735
 1638 GGAATTGAACATGAGCTCTGAGACAGAGACTCAAAAGAACCGGGGTGATATTTTGGACAG 1697
 1736 CACCATCCCTGATCTGTGCTGCTGCTGAGATGCTCCAGAGTCTCTTCTTCTTAAAC 1795
 1698 CACCATCCCTGATCTGTGCTGCTGCTGAGATGCTCCAGAGTCTCTTCTTCTTAAAC 1757
 1796 TTTCTCTCTCTCTCCAGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1855
 1758 TTTCTCTCTCTCTCCAGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1817
 1856 GTGATTTGTGCTTCT 1915
 1818 GTGATTTGTGCTTCT 1877
 1916 GCTGAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1975
 1878 GCTGAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1937
 1976 ACTTAGGAATGATCTTACAGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2035
 1938 ACTTAGGAATGATCTTACAGAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1997
 2036 GGCACCAAGATTTCTTCAATTAATTTTGAAGAGGCTGCTTATGATGATTAATTAAT 2095
 1998 GGCACCAAGATTTCTTCAATTAATTTTGAAGAGGCTGCTTATGATGATTAATTAAT 2057
 2096 GGTTTTGAATCTCATAGAAATTAATCAAGATGT 2130
 2058 GGTTTTGAATCTCATAGAAATTAATCAAGATGT 2092

RESULT 6
BC069501
DEFINITION Homo sapiens hypothetical protein FLJ31547, mRNA (cDNA clone MGC:97182 IMAGE:7262427), complete cds.
ACCESSION BC069501
VERSION BC069501.1 GI:46854682
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2047)
Strauberg, R.L., Feinberg, E.A., Grouse, J.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schley, G.D., Altschul, S.F., Zeeberg, B., Burow, K.H., Schaefer, C.P., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marinova, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.D., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richarde, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, J., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skelske, U., Small, D.B., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
human and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2047)
Director MGC Project.
Direct Submission
Submitted (29-APR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Baylor Human Genome Sequencing Center
cDNA Library Preparation: Baylor Human Genome Sequencing Center
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hui, S.W., Louised, H., Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaev, A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
Series: IRBR Plate: 4 Row: e Column: 1.
Location/Qualifiers
1. 2047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:97182 IMAGE:7262427"
/tissue_type="PCR rescued clones"
/clone_id="NIH_MGC_277"
/note="Vector: pPCR-Script Amp SK(+) with reversed insert"
1. 2047
/gene="FLJ31547"
/db_xref="locusID:221223"
121. 1848

/gene="FLJ31547"
/codon_start=1
/product="FLJ31547 protein"
/protein_id="AAH69501.1"
/db_xref="GI:46854683"
/db_xref="GI:46854682"
/translation="MSGWVHAGQILITWALVTLAAPKPSBAGPQNTLGIQKQVTVLGSFVNVFGLVFPAPPLSLRFTNPQPSFVMDNLRATSPNLCIONSEMLLDQHLKRVHFKFVSEBDCYLNTYAPADHDSGLPVLWFPFGAGKTSASIFDSSALAAVEDLVVVVYRGLRGFTTMDQHPAGNMAFKDOVAALSMVKNEPFGDPSVTTIFGSAQASIVSSSLISPMAGLPHKAIMESGVALITYLEAHYKEDQVAVHPCGNASPSSEALCLRTKPKELLTTSQKTSPTRVVDCAFPENELDLSQAFALPSITGVNHEGCGFLPKKAPRPLSSNSLALHLQNLHLPQYLHVAAEYFADKHSITRDSLDLIDGVFVFPVPLITARHHRGAPVYEFERHRCPEQEDRPVFKADHADRFLVFGAGFLKGIIVMEGATBEKLSRKMVMYATFATGPNQNDLSLWPAVNLQYOLQDLNLSGLQRLEKPEVDEWTSITIPILASDMLHSPLSLFPLSLQPFPPFCAP"

ORIGIN

Query Match 86.0%; Score 1855; DB 9; Length 2047;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	186	GGGCTTCTGCTGAAGGAGCCACAGAGAACACGAGCTGGATGATTCAGGGCAAGCAA	245
Db	193	GGGCTTCTGCTGAAGGAGCCACAGAGAACACGAGCTGGATGATTCAGGGCAAGCAA	252
Qy	246	GTCATCTGCTGGAGAACCTCTGCTGTAACGTGTCTCGAGTCCCTTTGCTGCT	305
Db	253	GTCATCTGCTGGAGAACCTCTGCTGTAACGTGTCTCGAGTCCCTTTGCTGCT	312
Qy	306	CCCCGCTGGAGATCCCTGCGATTGAAGAACCGGACGCTGATCGCCGGAATCTTG	365
Db	313	CCCCGCTGGAGATCCCTGCGATTGAAGAACCGGACGCTGATCGCCGGAATCTTG	372
Qy	366	CGAAGACCACTCTCTCACTTAATTTGTGCTCAGAACTCAGAGTGGCTCTTAAGAT	425
Db	373	CGAAGACCACTCTCTCACTTAATTTGTGCTCAGAACTCAGAGTGGCTCTTAAGAT	432
Qy	426	CAACAATGCTCAAGGTGATTAACCGGAATTCGGAGTGTCAAGAACTGCTCTACCTG	485
Db	433	CAACAATGCTCAAGGTGATTAACCGGAATTCGGAGTGTCAAGAACTGCTCTACCTG	492
Qy	486	AACATCTATGCGCTCTCCACGCGATACAGGCTCCAGCTCCGCTCTTGATGATGTC	545
Db	493	AACATCTATGCGCTCTCCACGCGATACAGGCTCCAGCTCCGCTCTTGATGATGTC	552
Qy	546	CCAGAGGTGCTTCAAGCTGAGCTCAGCTCAATCTTTGATGGTCCGCTGCTGCTC	605
Db	553	CCAGAGGTGCTTCAAGCTGAGCTCAGCTCAATCTTTGATGGTCCGCTGCTGCTC	612
Qy	606	TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	665
Db	613	TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	672
Qy	666	ACATGGGATCAGCATGCTCCGAGGAACTGAGGCTTCAAGAACAGAGTGGCTGCTGCTC	725
Db	673	ACATGGGATCAGCATGCTCCGAGGAACTGAGGCTTCAAGAACAGAGTGGCTGCTGCTC	732
Qy	726	TGGGTCACAGAAAGATGAGTCTTGGTGGGAGCCCAAGCTCTGTGACCATCTTTGGC	785
Db	733	TGGGTCACAGAAAGATGAGTCTTGGTGGGAGCCCAAGCTCTGTGACCATCTTTGGC	792
Qy	786	GAGTCGCGGAGGACATAGTCTTCTAGTCTTATAGTCTTCCATGAGCAAGGCTTA	845
Db	793	GAGTCGCGGAGGACATAGTCTTCTAGTCTTATAGTCTTCCATGAGCAAGGCTTA	852
Qy	846	TTCCCAAGAGCATATAGAGAGTGGGTCGATCATCTTCACTTGGAGGCCCATGAT	905
Db	853	TTCCCAAGAGCATATAGAGAGTGGGTCGATCATCTTCACTTGGAGGCCCATGAT	912
Qy	906	TATGAGAGAGTGAAGCTCAGAGTGGTTCACATTTCTGTGTAACATGCTGAC	965

Db 913 TATGAGAGAGTGAAGAGCTGACAGTGTGTCACATTTCTGTGTGAACAATGCGTCAGAC 972
 Qy 966 TCTGAGGCGCTGTGAGTGGCTGAGAGAGAAAACCTCCAGAGAGCTGTCAGCCCTCAGC 1025
 Db 973 TCTGAGGCGCTGTGAGTGGCTGAGAGAGAAAACCTCCAGAGAGCTGTCAGCCCTCAGC 1032
 Qy 1026 CAGAAAACAAAGCTCTTCACTGAGAGTGTGATGTGCTTCTTCTTAATAGAGCTCTTA 1085
 Db 1033 CAGAAAACAAAGCTCTTCACTGAGAGTGTGATGTGCTTCTTCTTAATAGAGCTCTTA 1092
 Qy 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATAACAC 1145
 Db 1093 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATAACAC 1152
 Qy 1146 GAGTGTGCTTCTGTGCTGCTAATGAAGAGGCTCCGAGATCCTCATGAGGCTCCCAACAG 1205
 Db 1153 GAGTGTGCTTCTGTGCTGCTAATGAAGAGGCTCCGAGATCCTCATGAGGCTCCCAACAG 1212
 Qy 1206 TCCCTTGCCCTCATCTGATACAAAACATCTGTCACATCCCGCTCAGTATTGACCTT 1265
 Db 1213 TCCCTTGCCCTCATCTGATACAAAACATCTGTCACATCCCGCTCAGTATTGACCTT 1272
 Qy 1266 GTGGCTAATGAATTAATCTTCATGACAGCACTCCCTGACTGAATCCGAGACAGTCTTCTG 1325
 Db 1273 GTGGCTAATGAATTAATCTTCATGACAGCACTCCCTGACTGAATCCGAGACAGTCTTCTG 1332
 Qy 1326 GACTTGCTTGAAGATGTGTTTGTGTGTCCTGTCATGATCAGAGTGTGATATCAACA 1385
 Db 1333 GACTTGCTTGAAGATGTGTTTGTGTGTCCTGTCATGATCAGAGTGTGATATCAACA 1392
 Qy 1386 GATGCTGTGTCACCTGTCTACTTCTTAATGATTTGCGACACGCGCTCAGTCTTGAAGAC 1445
 Db 1393 GATGCTGTGTCACCTGTCTACTTCTTAATGATTTGCGACACGCGCTCAGTCTTGAAGAC 1452
 Qy 1446 AGGAAAGCCGCTTTGTCTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCGTGTG 1505
 Db 1453 AGGAAAGCCGCTTTGTCTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCGTGTG 1512
 Qy 1506 GCGTTCCTGAAGGGGACATTTGTATGTTTGAAGAGCCAGAGAGAGAGAAAGTTACTG 1565
 Db 1513 GCGTTCCTGAAGGGGACATTTGTATGTTTGAAGAGCCAGAGAGAGAGAAAGTTACTG 1572
 Qy 1566 AGCCGGAAGATGATGAATTAATCTGAGTCTTGTGTCGAACCGGGAATCTTAATGGAAC 1625
 Db 1573 AGCCGGAAGATGATGAATTAATCTGAGTCTTGTGTCGAACCGGGAATCTTAATGGAAC 1632
 Qy 1626 GACCTGTCTCTGTGGCAGCTTAATTAATCTGATGAGCAGTACCTCCAGCTGGAATTGAAC 1685
 Db 1633 GACCTGTCTCTGTGGCAGCTTAATTAATCTGATGAGCAGTACCTCCAGCTGGAATTGAAC 1692
 Qy 1686 ATGAGCTTGGACAGAGATCAAGAAACCGCGGGTGGATTTTGAACGACACATCCCC 1745
 Db 1693 ATGAGCTTGGACAGAGATCAAGAAACCGCGGGTGGATTTTGAACGACACATCCCC 1752
 Qy 1746 CTGATCTCTGTGCTCCGACAGATGTCACAGTCTCTTCTTCTTAATCTTCCCTCTC 1805
 Db 1753 CTGATCTCTGTGCTCCGACAGATGTCACAGTCTCTTCTTCTTAACTTCTCTCTC 1812
 Qy 1806 CTCTCTCAGCCTTCTTCTTCTTCTTGTGTGCTCTTGAAGATTATCTTCTGTGATTTTG 1865
 Db 1813 CTCTCTCAGCCTTCTTCTTCTTGTGTGCTCTTGAAGATTATCTTCTGTGATTTTG 1872
 Qy 1866 TTTCCCTTCTCTCCATTAATTTCTCCGCAATCATTAAGTCTTCTTCTGAGCTCAGCTGC 1925
 Db 1873 TTTCCCTTCTCTCCATTAATTTCTCCGCAATCATTAAGTCTTCTTCTGAGCTCAGCTGC 1932
 Qy 1926 TTTCTAATGAGGATCTCTGCAAAACAAGCTTGTGCTGATATTTTAATGAGCTTAAGAT 1985
 Db 1933 TTTCTAATGAGGATCTCTGCAAAACAAGCTTGTGCTGATATTTTAATGAGCTTAAGAT 1992
 Qy 1986 GATCTTACAGATTTCTTTCAACATCAAAAGTGAATTTGTCTTGAAGGCA 2040
 Db 1993 GATCTTACAGATTTCTTTCAACATCAAAAGTGAATTTGTCTTGAAGGCA 2047

RESULT 7
 BC039073
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Kunarine, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Loulseg, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAX Plate: 51 Row: k Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 Location/Qualifiers
 source 1..2298
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5174937"
 /tissue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_id="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 ORIGIN
 Query Match 84.6%; Score 1824.6; DB 9; Length 2298;
 Best Local Similarity 93.5%; Pred No. 0;
 Matches 1969; Conservative 0; Mismatches 4; Indels 132; Gaps 1;
 Qy 186 GGGCTTGTGCTGAAGGGGCAAGAGAACACAGAGCTGGATGATTCAGGGCAAGCA 245
 Db 103 GGGCTTGTGCTGAAGGGGCAAGAGAACACAGAGCTGGATGATTCAGGGCAAGCA 162
 Qy 246 GTCACTGTGCTGGAAAGCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTGTGCT 305
 Db 163 GTCACTGTGCTGGAAAGCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTGTGCT 222
 Qy 306 CCCCCGCTGGATCCCTGCGATTTTGAACCCGAGAGCTGATCGCCCTGGATTAATTG 365
 Db 223 CCCCCGCTGGATCCCTGCGATTTTGAACCCGAGAGCTGATCGCCCTGGATTAATTG 282
 Qy 366 CGAAGCCACTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAT 425
 Db 283 CGAAGCCACTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAT 342

426 CAAGATGCTCAAGTGCATTAACCCGAAATTCGAGTGTCAAGAACTGCTCTACCTG 485
433 CAACACATGCTCAAGTGCATTAACCCGAAATTCGAGTGTCAAGAACTGCTCTACCTG 402
486 AACATCTATGCGCTGCGCCAGCAGATCAAGGCTCAAGGCTCCCGGCTTGGTGGTTC 545
403 AACATCTATGCGCTGCGCCAGCAGATCAAGGCTCCCGGCTTGGTGGTTC 462
546 CCGAGAGTGTGCTTCAAGACTGCTGAGCTCCATCTTTGATGAGGTCGCGCTGCTGCC 605
463 CCGAGAGTGTGCTTCAAGACTGCTGAGCTCCATCTTTGATGAGGTCGCGCTGCTGCC 522
606 TATGAGAGCTGTGCTGTTGATGTGTGTCAGATCCGCTAGAGAAATATTTGGTTCTTAC- 664
523 TATGAGAGCTGTGCTGTTGATGTGTGTCAGATCCGCTAGAGAAATATTTGGTTCTTACG 582
665 ----- 664
583 TGAATCTCTCAATGATGAGCATCCAGATCCCAAGGCGAGAGTGAAGTACCATTA 642
665 ----- 664
643 GACAGTGCAGCACTTAAGAAAGAGCGAGCTTCGCGGACATGAACAAACTGCGCC 702
665 -----CAATGGATTCAGCATGCTCCGCGGAACTGGGCTTCAAGACAGAGTG 713
703 TAATTAATTAGCAGATGGGATGAGCATGCTCGCGGAACTGGGCTTCAAGACAGAGTG 762
714 GCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 773
763 GCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822
774 ACCATCTTGGGAGTCCGCGGAGCATTAAGTGTCTTCAATTAATCTGCTCCAGT 833
823 ACCATCTTGGGAGTCCGCGGAGCATTAAGTGTCTTCAATTAATCTGCTCCAGT 882
834 GCGAAAGGCTTATTCACAAAGCATGAGAGTGGGAGCATCAATCCCTTACCTG 893
883 GCGAAAGGCTTATTCACAAAGCATGAGAGTGGGAGCATCAATCCCTTACCTG 942
894 GAGGCCATGATTAATGAGAAAGTGAAGCTGAGAGTGGGAGCATCAATCCCTTACCTG 953
943 GAGGCCATGATTAATGAGAAAGTGAAGCTGAGAGTGGGAGCATCAATCCCTTACCTG 1002
954 AATGCGTCAAGTCTGAGGCTGCTGAGTGCCTGAGGACAAAACCTCCAGAGAGCTG 1013
1003 AATGCGTCAAGTCTGAGGCTGCTGAGTGCCTGAGGACAAAACCTCCAGAGAGCTG 1062
1014 CTGACCTCAAGCAGAAACAAAGTCTTCACTGAGTGTGATGAGTGTCTTCTTCT 1073
1063 CTGACCTCAAGCAGAAACAAAGTCTTCACTGAGTGTGATGAGTGTCTTCTTCT 1122
1074 AATGAGCTTGAATCTATGCTCAGAAAGCATTTAAAGCAATCTTCCATCATGAGA 1133
1123 AATGAGCTTGAATCTATGCTCAGAAAGCATTTAAAGCAATCTTCCATCATGAGA 1182
1134 GTCAATTAACCAAGTGTGCTTCTGCTGCTTCAATGAAGGCTCTGAGATCTCAGT 1193
1183 GTCAATTAACCAAGTGTGCTTCTGCTGCTTCAATGAAGGCTCTGAGATCTCAGT 1242
1194 GGTCTCAACAAAGTCCCTGCTCCTCATCTGATACAAAACATCCGACATCCGCTCAG 1253
1243 GGTCTCAACAAAGTCCCTGCTCCTCATCTGATACAAAACATCCGACATCCGCTCAG 1302
1254 TATTTGCACTTGTGGCTAATGAATTAATTCATGACAAAGCTCCCTGATGAATCCGA 1313
1314 TATTTGCACTTGTGGCTAATGAATTAATTCATGACAAAGCTCCCTGATGAATCCGA 1362
1363 GACAGTCTTGTGAGCTTGTGAGATGTGTTCTTGTGTGCTCCTGACATGATCAGCT 1422

1374 CGATATCAGAGAGATGCTGTCACCTGTCTACTTCTATGAGTTTGGGACCGGCTGAG 1433
1423 CGATATCAGAGAGATGCTGTCACCTGTCTACTTCTATGAGTTTGGGACCGGCTGAG 1482
1434 TCGTTTGAAGACACAGAGCGGCTTTGTCAAGGCGACACAGCTGATGAATCGCTTT 1493
1483 TCGTTTGAAGACACAGAGCGGCTTTGTCAAGGCGACACAGCTGATGAATCGCTTT 1542
1494 GTGTTGCGTGTGCTCTCTGTAAGGGGACATTTGTTATGTTGAGAGGACACGAGAG 1553
1543 GTGTTGCGTGTGCTCTCTGTAAGGGGACATTTGTTATGTTGAGAGGACACGAGAG 1602
1554 GAGAGTTTACTGAGCGCGGAGATGATGAATTAATGAGGCTTACCTTGTCTGAAACGGG 1613
1603 GAGAGTTTACTGAGCGCGGAGATGATGAATTAATGAGGCTTACCTTGTCTGAAACGGG 1662
1614 CCTAATGAGGAGCGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1673
1663 CCTAATGAGGAGCGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1722
1674 CTGGAATTGAACATGAGGCTGCGACAGAGCTCAAGAACCGCGGAGTATTTTGGACC 1733
1723 CTGGAATTGAACATGAGGCTGCGACAGAGCTCAAGAACCGCGGAGTATTTTGGACC 1782
1734 AGACCAATCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793
1783 AGACCAATCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
1794 ACTTCCCT 1853
1843 ACTTCCCT 1902
1854 CTGGAATTTGGTGTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1913
1903 CTGGAATTTGGTGTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1962
1914 GAGCTCAGCTGCTTCT 1973
1963 GAGCTCAGCTGCTTCT 2022
1974 GAGCTCAGCTGCTTCT 2033
2023 GAGCTCAGCTGCTTCT 2082
2034 AAGGCAACAAATTTCTTCAATTAATTTGGAAGAGGCTGCTTATGATGATCAATA 2093
2083 AAGGCAACAAATTTCTTCAATTAATTTGGAAGAGGCTGCTTATGATGATCAATA 2142
2094 ATGATTTGTAATCTCATATGAATAAATCAGAAATGAATAATGAATAAATGAATAA 2153
2143 ATGATTTGTAATCTCATATGAATAAATCAGAAATGAATAATGAATAAATGAATAA 2202
2154 AAAAA 2158
2203 AAAAA 2207

RESULT 8
AR438765
LOCUS
DEFINITION
AR438765
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1746)
AUTHORS
Curtis, R.A.J. and Siles-Santiago, I.
53010, a human carboxylesterase family member and uses thereof
JOURNAL
Patent: us 6664091-A 3 16-DEC-2003;
FEATURES
Location/Qualifiers
1..1746

ORIGIN

/organism="unknown"
/mol_type="genomic DNA"

Query Match 80.9%; Score 1746; DB 6; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

96 ATGCCACAGGAGCTTACTCATCTGCTTCAACAATGATGCTTTTCTGATTTCTCAAGCCC 155
1 ATGCCACAGGAGCTTACTCATCTGCTTCAACAATGATGCTTTTCTGATTTCTCAAGCCC 60
156 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
216 ACCAGGCTGGGATGGAATTCAGGGCAAGATCTGTGCTGGAAAGCCCTGTGCTGTG 275
121 ACCAGGCTGGGATGGAATTCAGGGCAAGATCTGTGCTGGAAAGCCCTGTGCTGTG 180
276 AACGTTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTTACGAAAC 335
181 AACGTTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTTACGAAAC 240
336 CCGACGCTGCATCCGCTGGGATTAATTCGAGAAAGCCACTCTAATTTATTTGTC 395
241 CCGACGCTGCATCCGCTGGGATTAATTCGAGAAAGCCACTCTAATTTATTTGTC 300
396 CTCGAAACTCAGAGTGGTCTCTTAATCAACAATGCTCAAGTGCATTAACCGAAA 455
301 CTCGAAACTCAGAGTGGTCTCTTAATCAACAATGCTCAAGTGCATTAACCGAAA 360
456 TTCCGAGTGCAGAGTGGTCTCTTAATCAACAATGCTCAAGTGCATTAACCGAAA 515
361 TTCCGAGTGCAGAGTGGTCTCTTAATCAACAATGCTCAAGTGCATTAACCGAAA 420
516 GAGCTCAAGCTCCGCTCTGTTGTTGTTCCAGAGAGTGCCTTCAAGAAGCTCAGCC 575
421 GAGCTCAAGCTCCGCTCTGTTGTTGTTCCAGAGAGTGCCTTCAAGAAGCTCAGCC 480
576 TCCATCTTTGATGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 635
481 TCCATCTTTGATGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
636 TACCGGCTAGAGAAATTTGTTGTTCTTCAACAATGCTCAAGTGCATTAACCGAAA 695
541 TACCGGCTAGAGAAATTTGTTGTTCTTCAACAATGCTCAAGTGCATTAACCGAAA 600
696 GAGCTCAAGAGACAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
601 GAGCTCAAGAGACAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
756 GGGGACCCAGCTCTGTGACATCTTTTGGCGAGTCCGCGGAGCAATTAAGTCTTCTAG 815
661 GGGGACCCAGCTCTGTGACATCTTTTGGCGAGTCCGCGGAGCAATTAAGTCTTCTAG 720
816 CTTAATCTGTCTCCATGAGCCAAAGCTTATTCACAAAGCATCATGAGAGTGGGCTG 875
721 CTTAATCTGTCTCCATGAGCCAAAGCTTATTCACAAAGCATCATGAGAGTGGGCTG 780
876 GGCATCATCCCTTACCTGAGGCGCATGATTAAGAGAGAGAGAGAGAGAGAGAGAG 935
781 GGCATCATCCCTTACCTGAGGCGCATGATTAAGAGAGAGAGAGAGAGAGAGAGAG 840
936 GGCATCATCCCTTACCTGAGGCGCATGATTAAGAGAGAGAGAGAGAGAGAGAGAG 995
841 GGCATCATCCCTTACCTGAGGCGCATGATTAAGAGAGAGAGAGAGAGAGAGAGAG 900
996 AAACCTTCAAGAGAGTGTGACCTTCAAGCCGAGAAACAAATCTTTTCACTGAGTGT 1055
901 AAACCTTCAAGAGAGTGTGACCTTCAAGCCGAGAAACAAATCTTTTCACTGAGTGT 960
1056 GATGTGCTTTCTTCTTAATGAGCCTTAATGATTAATGCTCAGAAAGCATTTAAGCA 1115

961 GATGTGCTTTCTTCTTAATGAGCCTTAATGATTAATGCTCAGAAAGCATTTAAGCA 1020
1116 ATTCTTTCATCATATGAGAGTGAATAACAAGAGTGGCTTCCGCTGCTATGAAGAG 1175
1021 ATTCTTTCATCATATGAGAGTGAATAACAAGAGTGGCTTCCGCTGCTATGAAGAG 1080
1176 GCTCTGAGATCTCAGTGGCTCCAAAGATCCCTTGGCTCATCTGATACAAACATC 1235
1081 GCTCTGAGATCTCAGTGGCTCCAAAGATCCCTTGGCTCATCTGATACAAACATC 1140
1236 CTGACATCCGCTCAGTATTTGACCTTGTGCTAATGAATCTTCATGACAGAC 1295
1141 CTGACATCCGCTCAGTATTTGACCTTGTGCTAATGAATCTTCATGACAGAC 1200
1236 TCCCTGATGAAATCCGAGACAGTCTTGTGACCTTGTGAGAGTGTCTTGTGTC 1355
1201 TCCCTGATGAAATCCGAGACAGTCTTGTGACCTTGTGAGAGTGTCTTGTGTC 1260
1356 CCGACATGATCAAGCTGATATCAAGAGATGCTGTCAGCTGTCTATGAG 1415
1261 CCGACATGATCAAGCTGATATCAAGAGATGCTGTCAGCTGTCTATGAG 1320
1416 TTTCGCAACCGGCTCAGTCTTTGAGACAGAAAGCCGCTTTTGTCAAGCCGACAC 1475
1321 TTTCGCAACCGGCTCAGTCTTTGAGACAGAAAGCCGCTTTTGTCAAGCCGACAC 1380
1476 GCTGATGAATCCGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1535
1381 GCTGATGAATCCGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
1536 GAGGAGCCACGAGAGAGAGAGAGTACTGAGCCGAGAGATGATGAATTAATGCTACC 1595
1441 GAGGAGCCACGAGAGAGAGAGAGTACTGAGCCGAGAGATGATGAATTAATGCTACC 1500
1596 TTTCGCAACCGGAAATCTTAATGAGACAGACCTGTCTGTGTCAGCTTAATGCTG 1655
1501 TTTCGCAACCGGAAATCTTAATGAGACAGACCTGTCTGTGTCAGCTTAATGCTG 1560
1656 ACTGAGCATACCTTCAAGCTGATCTTGAACATGAGCTTCCGACAGAGACTCAAAAGAC 1715
1561 ACTGAGCATACCTTCAAGCTGATCTTGAACATGAGCTTCCGACAGAGACTCAAAAGAC 1620
1716 GGGGAGATTTTGGACAGACCAATCCCGTGAATCTGTGCTGCTCCGACATGCTCCAC 1775
1621 GGGGAGATTTTGGACAGACCAATCCCGTGAATCTGTGCTGCTCCGACATGCTCCAC 1680
1776 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
1681 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1836 CCTTGA 1841
1741 CCTTGA 1746

RESULT 9
AX600126 1746 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 3 from Patent WO0250256.
ACCESSION AX600126
VERSION AX600126.1 GI:28400192
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Curtis, R.A. and Silos-Santiago, I.
53010, a novel human carboxylesterase family member and uses
thereof
Patent: WO 0250256-A 3 27-JUN-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)

FEATURES
source
Location/Qualifiers
1. 1746
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 80.9%; Score 1746; DB 6; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 ATGCCACAGGAGCTTACTTCAATCTGCTTCACAATAGTGTCTTTTCTGATTTCTCCAGCCC 155
Db 1 ATGCCACAGGAGCTTACTTCAATCTGCTTCACAATAGTGTCTTTTCTGATTTCTCCAGCCC 60
QY 156 CTGTGGGAGACACAGACAGTGGGGAAAAATCGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
Db 61 CTGTGGGAGACACAGACAGTGGGGAAAAATCGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
QY 216 ACCAGGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTG 275
Db 121 ACCAGGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTG 180
QY 276 AAGTGTCTCTGGAGTCCCTTTTGTCTCTCCCGCTGGATCCCTGGATTTTACGAAC 335
Db 181 AAGTGTCTCTGGAGTCCCTTTTGTCTCTCCCGCTGGATCCCTGGATTTTACGAAC 240
QY 336 CCGCAGCTGCAATCGGCTGGGATTAATTGCGAGAAACCACTCTCAACCTTAATTTGTGC 395
Db 241 CCGCAGCTGCAATCGGCTGGGATTAATTGCGAGAAACCACTCTCAACCTTAATTTGTGC 300
QY 396 CTCCAGAACTCAGAGTGGTGTCTTATGATCAACATGCTCAAGGTGATTAACCGAAA 455
Db 301 CTCCAGAACTCAGAGTGGTGTCTTATGATCAACATGCTCAAGGTGATTAACCGAAA 360
QY 456 TTCCGAGTCTCAGAGTGGTGTCTTATGATCAACATGCTCAAGGTGATTAACCGAAA 515
Db 361 TTCCGAGTCTCAGAGTGGTGTCTTATGATCAACATGCTCAAGGTGATTAACCGAAA 420
QY 516 GGGTCCAAAGCTCCCGCTTGTGTGTGTCCAGAGAGTGCCTTCAAGATCGGCTCAACC 575
Db 421 GGGTCCAAAGCTCCCGCTTGTGTGTGTCCAGAGAGTGCCTTCAAGATCGGCTCAACC 480
QY 576 TCCATCTTTGATGAGTCCCGCTGCTGCTATGAGAACGTGTGTGTGTGTGTGTGTGTGT 635
Db 481 TCCATCTTTGATGAGTCCCGCTGCTGCTATGAGAACGTGTGTGTGTGTGTGTGTGTGT 540
QY 636 TACCGGCTAGGAATATTTGT 695
Db 541 TACCGGCTAGGAATATTTGT 600
QY 696 GCTTCAAGAACACAGTGGT 755
Db 601 GCTTCAAGAACACAGTGGT 660
QY 756 GGGGACCCCAAGCTCTGTGACATCTTTGGAGAGTCCGGGGAGCAATAGTGTGTGTGTGT 815
Db 661 GGGGACCCCAAGCTCTGTGACATCTTTGGAGAGTCCGGGGAGCAATAGTGTGTGTGTGT 720
QY 816 CTTAATCTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTG 875
Db 721 CTTAATCTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTG 780
QY 876 GGCATCATCCCTTAACCTGAGAGGCCCATGATTTATGAGAAAGTGTGTGTGTGTGTGT 935
Db 781 GGCATCATCCCTTAACCTGAGAGGCCCATGATTTATGAGAAAGTGTGTGTGTGTGTGT 840
QY 936 GCACATTTCTGT 995
Db 841 GCACATTTCTGT 900
QY 996 AAACCTCTCAAGAGT 1055

Db 901 AAACCTCTCAAGAGT 960
QY 1056 GATGGGCTTTCTCTTCTTAATGAGCTCTAATCTTATGTCTCAGAAAGCATTTAAAGCA 1115
Db 961 GATGGGCTTTCTCTTCTTAATGAGCTCTAATCTTATGTCTCAGAAAGCATTTAAAGCA 1020
QY 1116 ATTCTTCATCATCGAGTCAATTAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1175
Db 1021 ATTCTTCATCATCGAGTCAATTAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1176 GCTCTGTGATTCCTGATGT 1235
Db 1081 GCTCTGTGATTCCTGATGT 1140
QY 1236 CTGCACATCCCGCTCAGTATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1295
Db 1141 CTGCACATCCCGCTCAGTATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 1296 TCCCTGACTGAATCCGAGACAGTCTTGTGACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1355
Db 1201 TCCCTGACTGAATCCGAGACAGTCTTGTGACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
QY 1356 CCTGACATGATCAGAGTCCATATACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1415
Db 1261 CCTGACATGATCAGAGTCCATATACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
QY 1416 TTTGGGACCGGCTCAGTGTCTTTGAAGACAGAAAGCGGCTTTTGTCAAAAGCGAAC 1475
Db 1321 TTTGGGACCGGCTCAGTGTCTTTGAAGACAGAAAGCGGCTTTTGTCAAAAGCGAAC 1380
QY 1476 GCTGATGAATCCGCTTTGT 1535
Db 1381 GCTGATGAATCCGCTTTGT 1440
QY 1536 GAAGAGCCACAGAGAGAGAGAGTACTGAGCCGAGAGATGATGAATATCTGGGTATCC 1595
Db 1441 GAAGAGCCACAGAGAGAGAGAGTACTGAGCCGAGAGATGATGAATATCTGGGTATCC 1500
QY 1596 TTTGTCTGAAACCGGGAATCTTAATGAGAAAGACCTGTCTGTGTGTGTGTGTGTGTGT 1655
Db 1501 TTTGTCTGAAACCGGGAATCTTAATGAGAAAGACCTGTCTGTGTGTGTGTGTGTGTGT 1560
QY 1656 ACTGAGAGATCTCCAGCTGAGCTTGAACATGAGCTGTGACAGAGATCTCAAGAACCG 1715
Db 1561 ACTGAGAGATCTCCAGCTGAGCTTGAACATGAGCTGTGACAGAGATCTCAAGAACCG 1620
QY 1716 CCGGTGATTTTGTGAACAGACCATCCCTGATCTGTGTGTGTGTGTGTGTGTGTGTGT 1775
Db 1621 CCGGTGATTTTGTGAACAGACCATCCCTGATCTGTGTGTGTGTGTGTGTGTGTGTGT 1680
QY 1776 AGTCCCTTTCTTCTTAATCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
Db 1681 AGTCCCTTTCTTCTTAATCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY 1836 CCTTGA 1841
Db 1741 CCTTGA 1746

RESULT 10
AX714052
LOCUS
DEFINITION
Sequence 736 from Patent EP1293569.
AX714052
ACCESSION
AX714052.1 GI:29888980
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
1 Isegaki, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,

Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
Masuko, Y.
Full-length cDNAs
Patent: EP 1293569-A 736 19-MAR-2003;
JOURNAL
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)

FEATURES
Source
1. 1962
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 73.2%; Score 1579.4; DB 6; Length 1962;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;
186 GGGCCCTTGTCTGAAGGGCCACAGAGAACACACAGCTGGAGTTCAGAGGCAAGCA 245
222 GGGCCCTTGTCTGAAGGGCCACAGAGAACACACAGCTGGAGTTCAGAGGCAAGCA 281
246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTGAGTCCCTTTGCTGCT 305
282 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTGAGTCCCTTTGCTGCT 341
306 CCCCCGTGGAGATCCCTGGAATTTACGAACCCGACGCTGCATCGCCCTGGAGTAACTTG 365
342 CCCCCGTGGAGATCCCTGGAATTTACGAACCCGACGCTGCATCGCCCTGGAGTAACTTG 401
366 CGAGAGCCACCTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAGT 425
402 CGAGAGCCACCTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAGT 461
426 CAACCATGCTCAAGTGTGATTAACCCGAATTCGAGAGTGTCAAGAGCTGCTCTTAAGT 485
462 CAACCATGCTCAAGTGTGATTAACCCGAATTCGAGAGTGTCAAGAGCTGCTCTTAAGT 521
486 AACATCTATGCGCTGCGCCAGCGAGTACAGAGCTCAAGAGCTCCCGCTTGGTGTGTTG 545
522 AACATCTATGCGCTGCGCCAGCGAGTACAGAGCTCAAGAGCTCCCGCTTGGTGTGTTG 581
546 CCAGAGAGTGCCTTCAAGACTGCTCAAGCTTCATCTTTGATGGTCCGCTGCTGCTGCT 605
582 CCAGAGAGTGCCTTCAAGACTGCTCAAGCTTCATCTTTGATGGTCCGCTGCTGCTGCT 641
606 TATGAGAGCTGCTGCTGCTGCTGCTCAAGTACCGGCTTACGAATTTTGGTCTTCAAC 665
642 TATGAGAGCTGCTGCTGCTGCTGCTCAAGTACCGGCTTACGAATTTTGGTCTTCAAC 701
666 ACATGAGATCAGAGCTGCTGCTGCTGCTGCTCAAGAGCAAGAGTGGCTGCTGCTGCT 725
702 ACATGAGATCAGAGCTGCTGCTGCTGCTGCTCAAGAGCAAGAGTGGCTGCTGCTGCT 761
726 TGGGTTCAGAGAGACATGAGTCTTCCGTGGGAGACCCAGCTCTGTGACATCTTTGGC 785
762 TGGGTTCAGAGAGACATGAGTCTTCCGTGGGAGACCCAGCTCTGTGACATCTTTGGC 821
786 GAGTCCGCGGAGACATGAGTCTTCTAGTCTTAACTGCTCTCCATGGCCAAAGGCTTA 845
822 GAGTCCGCGGAGACATGAGTCTTCTAGTCTTAACTGCTCTCCATGGCCAAAGGCTTA 881
846 TTTCCCAAGAGCATATGAGAGAGTGGGGGACATACCCCTTACCTGAGGCCCATAT 905
882 TTTCCCAAGAGCATATGAGAGAGTGGGGGACATACCCCTTACCTGAGGCCCATAT 941
906 TATGAGAGAGTGAAGACCTGAGAGTGGTTCACATTTCTGTGTAAACATGCTCAAGC 965
942 TATGAGAGAGTGAAGACCTGAGAGTGGTTCACATTTCTGTGTAAACATGCTCAAGC 1001
966 TCTGAGGCCCTGCTGAGTGTCTGAGAGCAAAACCTTCAGAGAGCTGCTGACCTCAAC 1025
1002 TCTGAGGCCCTGCTGAGTGTCTGAGAGCAAAACCTTCAGAGAGCTGCTGACCTCAAC 1061

1026 CAGAAAACAAAGTCTTTCACCTGAGTGTGATGAGTGTCTTTCTTCTTAATGAGCTCTTA 1085
1062 CAGAAAACAAAGTCTTTCACCTGAGTGTGATGAGTGTCTTTCTTCTTAATGAGCTCTTA 1121
1086 GATCTATTGTCTCAGAAAACATTTAAAGCATTTCTTCAATCATGTGAGTCAATTAACAC 1145
1122 GATCTATTGTCTCAGAAAACATTTAAACAAATTCCTTCAATCATGTGAGTCAATTAACAC 1181
1146 GAGTGTGCTTCCCTGCTGCTATGAGAGAGGCTCCTGAGATCCTGAGTGTCTCAACAG 1205
1182 GAGTGTGCTTCCCTGCTGCTATGAGAGAGGCTCCTGAGATCCTGAGTGTCTCAACAG 1241
1206 TCCCTTGGCCCTTCATCTGATACAAAACATCTGACATCCGCTCAGATATTGACCTT 1265
1242 TCCCTTGGCCCTTCATCTGATACAAAACATCTGACATCCGCTCAGATATTGACCTT 1301
1266 GTGGCTAATGAATTAATTCATGACAGACATCTCCCTGACTGAATTCGAGACAGTCTTCTG 1325
1302 GTGGCTAATGAATTAATTCATGACAGACATCTCCCTGACTGAATTCGAGACAGTCTTCTG 1361
1326 GACTGTGCTGAGATGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1385
1362 GACTGTGCTGAGATGATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1421
1386 GATGCTGTGACACTGTCTACTTATGATGATTTCCGACACCGGCTCAGTGTCTTGAAGC 1445
1422 ----- 1421
1446 ACGAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTGTCCGTGCT 1505
1422 ----- 1421
1506 GCCTTCTGAAGGGGGACATTTGTTATGTTGGAAGAGCCAGGAGAGAGAGTTACTG 1565
1422 ----- 1421
1422 ----- 1421
1566 AGCCGGAAGATGATGAATTAATCTGAGCTACTTGTGCTGCAACCGGAACTCTAATGGAAAC 1625
1452 AGCCGGAAGATGATGAATTAATCTGAGCTACTTGTGCTGCAACCGGAACTCTAATGGAAAC 1511
1626 GACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1685
1512 GACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1571
1686 ATGAGCTCTGGAAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1745
1572 ATGAGCTCTGGAAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1631
1746 CTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1805
1632 CTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1691
1806 CTCTCAGAGCTTCTTTTCTTTTCTTTGCTCTTGAAGAGTTATCTTCTGATTTTGG 1865
1692 CTCTCAGAGCTTCTTTTCTTTTCTTTGCTCTTGAAGAGTTATCTTCTGATTTTGG 1751
1866 TTTCCCTTCTCTCTCCCAATTTTCTCCCGCAATCAATGCTTCTTCTGAGCTCAAGCTG 1925
1752 TTTCCCTTCTCTCTCCCAATTTTCTCCCGCAATCAATGCTTCTTCTGAGCTCAAGCTG 1811
1926 TTTCTATGAGGATCTTTCAGAAAACAGAGCTGCTTTCGATATTTTATGAGCTTGAAGT 1985
1812 TTTCTATGAGGATCTTTCAGAAAACAGAGCTGCTTTCGATATTTTATGAGCTTGAAGT 1871
1986 GATCTTTCAGAAATCTTTTCAACATCAAAAAGTGAATTTTGTCTGGAAGGCAACAA 2045
1872 GATCTTTCAGAAATCTTTTCAACATCAAAAAGTGAATTTTGTCTGGAAGGCAACAA 1931
2046 TTTCTTCAATTAATTTTGAAGAGGCTGCTG 2076
1932 TTTCTTCAATTAATTTTGAAGAGGCTGCTG 1962

RESULT 11
AK056109
LOCUS
DEFINITION Homo sapiens cDNA FLJ31547 f1s, clone NT2R12001010, weakly similar to FAT2 ACY1-COA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
ACCESSION AK056109
VERSION AK056109.1 GI:16551422
KEYWORDS c1igo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Oka,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Makamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shitatori,A., Sudo,H., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ichibashi,T., Yamashita,H., Hirakawa,K., Fujimori,K., Tanai,H., Kimura,M., Watanabe,M., Yoshida,M., Chiba,Y., Ichida,S., Ono,Y., Takiguchi,S., Watanabe,S., Hara,H., Tamase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Muraashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuma,S., Yoshioka,Y., Matsumura,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terahashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mitoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiroo,M., Ohmori,Y., Kawabara,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togaishi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Maehuo,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isegai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

JOURNAL
PUBMED
REFERENCES
AUTHORS
2
Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T., Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuma,N., Sato,K., Tanikawa,M., Yamazaki,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Makamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Oshima,A., Sugiyama,A., Kanehori,K., Takahashi-Fujii,A., Nishikawa,K., Maehuo,Y., Nagai,K. and Isegai,T.
NEBO human cDNA sequencing project
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
COMMENT
3 (bases 1 to 1962)
Isegai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isegai, Helix Research Institute, Genomics Laboratory, 1532-3 Yena, Kisarazu, Chiba 222-0812, Japan (E-mail:genomics@hi.co.jp, Tel:81-338-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
Location/Qualifiers

source
1..1962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2R12001010"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_1ib="NT2R12"
/note="cloning vector: PMB18SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"
150..1727
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71094.1"
/db_xref="GI:16551423"
/translation="MSGNWYRPGQILITWAIWLAAPTKGPSAGPQNTRLNIGQKQVTVGSPVVPVVPFGVPPAPBPGLSLFTNPQASPDNDLREATSYENLCLNSEWLLDQMLKVNHPKEVSDCYLNTIYAPHADTGSKLVLVWFPQGAFTGSAIPIQGSALAAVEVLVAVVQVSRGIFGFTTWDQAHGAWAFQDQVALSWQKNIFFGQPS SVTFGSAASIVSLILSPMAKGLFHKAIMSGVALIPLYLAHDYKESDEDVAVH FCGNNADSSEALIRCTRTSKELTLOSKTSFTRVODGAFPPNEDLDSOKAKA IPSIIIVNHECGFLPMEKAPETLGSNSKSLAHLIIONTIHTI PPOYIHLVAVYERHD KSLITEIRDSLDLIDGVFPVPAITARVHREBGADEEKLRSKMKRTATPATRGN PNGNDLSIMPAVNTBQYIOLDDIMSLGQRLEKRPVDFWTSTIPLTILSASDMLSPLS SUTFLSLQEPFFPCAP"

ORIGIN
Query Match 73.2%; Score 1579.4; DB 9; Length 1962;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

186 GGGCTTCTGCTGAAGGGCCACAGAGACACCAAGCTGGATGATTCAGGGCAACAA 245
222 GGGCTTCTGCTGAAGGGCCACAGAGACACCAAGCTGGATGATTCAGGGCAACAA 281
246 GTCACTGTGCTGGAAGACCTGTGCTGGAAGCTGTGCTGGAAGCTGTGCTGCT 305
282 GTCACTGTGCTGGAAGACCTGTGCTGGAAGCTGTGCTGGAAGCTGTGCTGCT 341
306 CCCCCGCTGGATCCTGCGATTTCAGAACCCGACCTGCAATCGCCCTGGATTAATTG 365
342 CCCCCGCTGGATCCTGCGATTTCAGAACCCGACCTGCAATCGCCCTGGATTAATTG 401
366 CGAAGAGCCACCTTCAACCTTAATTTGCTTCCAGAACTCAAGATGCTGCTTTAGAT 425
402 CGAAGAGCCACCTTCAACCTTAATTTGCTTCCAGAACTCAAGATGCTGCTTTAGAT 461
426 CAACACATGCTCAAGGTGATTAACCGAATTCGAGTGTGAGAGACTGCTTACCGT 485
462 CAACACATGCTCAAGGTGATTAACCGAATTCGAGTGTGAGAGACTGCTTACCGT 521
486 AACATCTAAGCCCTGCGCCCAAGCCGATACAGGCTCCAAAGCTCCCGTCTGGTGTGCTT 545
522 AACATCTAAGCCCTGCGCCCAAGCCGATACAGGCTCCAAAGCTCCCGTCTGGTGTGCTT 581
546 CCAGAGAGTCCCTTCAAGACTGCTCAGCTTCATCTTGAATGGTTCGCGCTGGCTGCC 605
582 CCAGAGAGTCCCTTCAAGACTGCTCAGCTTCATCTTGAATGGTTCGCGCTGGCTGCC 641
606 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
642 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
666 ACATGAGATCAGCATGCTCCCGGGAACCTGAGGCTTCAAGACCAAGTGGCTGCTGCTGCT 725
702 ACATGAGATCAGCATGCTCCCGGGAACCTGAGGCTTCAAGACCAAGTGGCTGCTGCTGCT 761
726 TGGGTCCAGAAACATGAGATCTTTCGCTGGGAGACCCAGACTGTGACCATCTTTGGC 785
762 TGGGTCCAGAAACATGAGATCTTTCGCTGGGAGACCCAGACTGTGACCATCTTTGGC 821
786 GAGTCCGCGGAGCATTAAGTGTTCCTTAAGTCTTAACTGCTCCCATGGCCAAAGGCTTA 845

Db 822 GAGTCGCGGAGGACCAATAGTGTCTTACTCTTAATCTGCTCCATGCGCAAGGCTTA 881
 Qy 846 TTTCACCAAAAGCCATCATGAGAGAGTGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGAT 905
 Db 882 TTTCACCAAAAGCCATCATGAGAGAGTGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGAT 941
 Qy 906 TATGAGAAGATGAGGACCTGAGAGGTGTGACATTTCTGTGTAAACAATGCGTCAAG 965
 Db 942 TATGAGAAGATGAGGACCTGAGAGGTGTGACATTTCTGTGTAAACAATGCGTCAAG 1001
 Qy 966 TCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGTGTGACCTCTGAC 1025
 Db 1002 TCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGTGTGACCTCTGAC 1061
 Qy 1026 CAGAAAACAAATCTTTCACTGAGAGGTGTGATGTGTGTCTTTCTTCTTAATAGGCTCTA 1085
 Db 1062 CAGAAAACAAATCTTTCACTGAGAGGTGTGATGTGTGTCTTTCTTCTTAATAGGCTCTA 1121
 Qy 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACAC 1145
 Db 1122 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACAC 1181
 Qy 1146 GAGTGTGGCTTCTGCTGCTCTAATGAAGAGGCTCTGAGATCTCAAGTCTCAACAAAG 1205
 Db 1182 GAGTGTGGCTTCTGCTGCTCTAATGAAGAGGCTCTGAGATCTCAAGTCTCAACAAAG 1241
 Qy 1206 TCCCTTGGCCCTCATCTGATACAAACATCTCTGACATCCCGCTCAAGTATTTGACCTT 1265
 Db 1242 TCCCTTGGCCCTCATCTGATACAAACATCTCTGACATCCCGCTCAAGTATTTGACCTT 1301
 Qy 1266 GTGGGCTAATGAATCTCTCATGACAAAGCACTCCCTGACTGAATCCGAGACAGTCTCTG 1325
 Db 1302 GTGGGCTAATGAATCTCTCATGACAAAGCACTCCCTGACTGAATCCGAGACAGTCTCTG 1361
 Qy 1326 GACTTGTGTGAGATGTGTCTTTGTGTGTCTCTGACCTGACCTGATCAAGCTGATATACAGA 1385
 Db 1362 GACTTGTGTGAGATGTGTCTTTGTGTGTCTCTGACCTGACCTGATCAAGCTGATATACAGA 1421
 Qy 1386 GATGCTGTGACACTGTCTTACTTATGATTTGCGACCGGCTCAAGTCTTTGAAGAC 1445
 Db 1422 ----- 1421
 Qy 1446 ACGAAGCGGCTTTGTCAAGCCGACACGCGTGAATGAAGTCCGCTTTGTGTGCGTGT 1505
 Db 1422 ----- 1421
 Qy 1506 GCCTTCTGAGAGGGGACATTTGTATGTTGGAAGAGCCACGAGAGGAGAGAGTACTG 1565
 Db 1422 -----GAAAGAGCCACGAGAGGAGAGAGTACTG 1451
 Qy 1566 AGCCGGAAGATGATGAATATCTGGGCTACTTTGTCTGAAACCGGGAATCTTAATGGGAAC 1625
 Db 1452 AGCCGGAAGATGATGAATATCTGGGCTACTTTGTCTGAAACCGGGAATCTTAATGGGAAC 1511
 Qy 1626 GACCTGTCTGTGTGCGAGCTTATATCTGACTGAGACATACCTCCAGCTGAGCTTGAAC 1685
 Db 1512 GACCTGTCTGTGTGCGAGCTTATATCTGACTGAGACATACCTCCAGCTGAGCTTGAAC 1571
 Qy 1686 ATGAGCCTCGACAGAGACTCAAAAGAACCGGCGTGTGATTTTGTGACCGACCAATCCCC 1745
 Db 1572 ATGAGCCTCGACAGAGACTCAAAAGAACCGGCGTGTGATTTTGTGACCGACCAATCCCC 1631
 Qy 1746 CTGATCTGTCTGTCTCGACAGTCTCAAGTCTCTTTCTTCTTAACCTTCTCTCT 1805
 Db 1632 CTGATCTGTCTGTCTCGACAGTCTCAAGTCTCTTTCTTCTTAACCTTCTCTCT 1691
 Qy 1806 CTCCTCCAGCCTTTCTTTTCTTTGTGTCTCTGAGAGTATCTTTCTGAGCTTGAAGTTGG 1865
 Db 1632 CTCCTCCAGCCTTTCTTTTCTTTGTGTGTCTCTGAGAGTATCTTTCTGAGCTTGAAGTTGG 1751
 Qy 1866 TTTCCCTTCTCTCCCATTAATTTCTCCGCAATCATTAAGCTTCTTTGAGCTGACGCTGC 1925

Db 1752 TTTCCCTTCTCTCCCATTAATTTCTCCGCAATCATTAAGCTTCTTTGAGCTGACGCTGC 1811
 Qy 1926 TTTCTATGGGAGATCTCTTGCAAAAGAGCTGCTTTGCTGATATTTATGAGACTTAGAAT 1985
 Db 1812 TTTCTATGGGAGATCTCTTGCAAAAGAGCTGCTTTGCTGATATTTATGAGACTTAGAAT 1871
 Qy 1986 GATCCTTACAGAAATCTTTTCAACATCAAAAGTCAATTTGTCTTGAAGGCAACAAG 2045
 Db 1872 GATCCTTACAGAAATCTTTTCAACATCAAAAGTCAATTTGTCTTGAAGGCAACAAG 1931
 Qy 2046 TTTCTTCAATTAATTTGGAAGAGGCTGACC 2076
 Db 1932 TTTCTTCAATTAATTTGGAAGAGGCTGACC 1962

RESULT 12

BC069548

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

REFERENCES

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLES

1 (bases 1 to 1896)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stabileton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshyuk, S.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McMan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalski, U., Smalins, D.E.,
 Schermer, A., Schein, J.E., Jones, S.J., and Matra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1896)
 Strausberg, R.
 Direct Submission
 Submitted (29-Apr-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-rt@mail.nih.gov
 Tissue Procurement: Baylor Human Genome Sequencing Center
 cDNA Library Preparation: Baylor Human Genome Sequencing Center
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louieged, H.,
 Kowis, C.R., Sneed, A.D., Martin, R.G., Muzny, D.M., Navevati,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/INL ac: <http://image.lnl.gov>
Series: IRBR Plate: 4 Row: f Column: 1.
Location/Qualifiers
1. 1896

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:97193 IMAGE:7262439"
/issue_type="PCR rescued clones"
/clone_id="NH MGC 244"
/note="Vector: pPCR-script Amp SK(+)"
1. 1896
/gene="FLJ31547"
/db_xref="LocusID:221223"
120. 1697
/gene="FLJ31547"

CDS

/codon_start=1
/product="hypothetical protein FLJ31547"
/protein_id="AAH69548.1"
/db_xref="GI:46854431"
/translation="MSGNVVHPGQILIMAVILAAPTKPSAAGPQRRLGWIQCKQ
VTVLGSPPVNVPLGVPAPAPPLGSLRFTNPQAPSWDLREASTIPNLCLQNSBML
LQOHMLKHYVPRKGVSEDCILNITAPAHADTGSKLPLVWPFGAFKGSASIFDGS
ALAAIVDVLVVVQYRLGIFGFTWDQHPQNMWAKDQVALSWOKNIEFGDPS
SVTVFESAGALSVSLIISPMACKLPHRAIMESGVALIPLYEADHYKESDLQYVAH
FGNNASDEBALRCLRTKRSKEILTLISQKTSFRTVGVGAFPPNEPDLISOKRFA
ISLIGVNHCEGFLPMKEAPEILSGSKSLALHLIONLIHPOLYLIHIVANEFHD
KSLIETIRSDSLDLGDFPVVPLITATGATVREKATBEKLSLRKMKRWATFATGN
PQNDSLISLPAYNLTPQYIQLDLNSLGRLEKPEVDFWTSTIPILISADMLHSPLS
SLTPSLDLPFFPCAP"

ORIGIN

Query Match 71.5%; Score 1543.4; DB 9; Length 1896;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1704; Conservative 0; Mismatches 1; Indels 150; Gaps 1;
QY 186 GGGCCCTTCGCTGAAAGGCGCAAGAGAAACACAGGCTGGATGATTCAGGGCAAGCAA 245
DB 192 GGGCCCTTCGCTGAAAGGCGCAAGAGAAACACAGGCTGGATGATTCAGGGCAAGCAA 251
QY 246 GTCACTGCTGCTGGAGAGCCCTGCTGCTGAAAGCTTCCTCGAGTCCCTTGGCTCT 305
DB 252 GTCACTGCTGCTGGAGAGCCCTGCTGCTGAAAGCTTCCTCGAGTCCCTTGGCTCT 311
QY 306 CCCCCGCTGGAGTCCCTGCGATTTACGAACCCGCAAGCTGATCGCCCTGGGATTAATTG 365
DB 312 CCCCCGCTGGAGTCCCTGCGATTTACGAACCCGCAAGCTGATCGCCCTGGGATTAATTG 371
QY 366 CGAAGAGCACCTCTACCTTAATTTGCTCCGAACTCAGAGTGGCTCTTAAAT 425
DB 372 CGAAGAGCACCTCTACCTTAATTTGCTCCGAACTCAGAGTGGCTCTTAAAT 431
QY 426 CAACATGCTCAAGTGCATTTACCCGAAATTCGAGTGCAGAGATGCTCTTACCTG 485
DB 432 CAACATGCTCAAGTGCATTTACCCGAAATTCGAGTGCAGAGATGCTCTTACCTG 491
QY 486 AACATCTATGGCTGCTGCCAGCGGATACAGGCTTCAAGCTCCCGCTTGGTGGTTC 545
DB 492 AACATCTATGGCTGCTGCCAGCGGATACAGGCTTCAAGCTCCCGCTTGGTGGTTC 551
QY 546 CCAGAGAGTGCCTTCAAGATGAGCTCAGGCTCAATCTTTGATGGGTCCGCTGGCTGCC 605
DB 552 CCAGAGAGTGCCTTCAAGATGAGCTCAGGCTCAATCTTTGATGGGTCCGCTGGCTGCC 611
QY 606 TATAGAGAGCTGCTGCTGGTGGTGCATGACCGGTAGAGAAATTTGGTTCTTACG 665
DB 612 TATAGAGAGCTGCTGCTGGTGGTGCATGACCGGTAGAGAAATTTGGTTCTTACG 671
QY 666 ACATGGGATCAGATGCTCCGGGGAACTGGGCTTCAAGACCAAGTGGCTGCTGTGCC 725
DB 672 ACATGGGATCAGATGCTCCGGGGAACTGGGCTTCAAGACCAAGTGGCTGCTGTGCC 731

QY 726 TGGGTCCAGAAAGATCGAGTTCTTCGATGGGAGCCCCAGCTGTGACCATCTTTGGC 785
DB 732 TGGGTCCAGAAAGATCGAGTTCTTCGATGGGAGCCCCAGCTGTGACCATCTTTGGC 791
QY 786 GAGTCCGGGGAGCCATAGTCTTTCTAGTCTTATCTGTCTCCCATGGCCAAAGCTTAA 845
DB 792 GAGTCCGGGGAGCCATAGTCTTTCTAGTCTTATCTGTCTCCCATGGCCAAAGCTTAA 851
QY 846 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGAT 905
DB 852 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGAT 911
QY 906 TATGAGAGAGTGAAGACCTTGAGGTGTTGCACATTTCTGTGTAAACATGCTCGACAG 965
DB 912 TATGAGAGAGTGAAGACCTTGAGGTGTTGCACATTTCTGTGTAAACATGCTCGACAG 971
QY 966 TCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCCCTCCAAAGAGCTGTGACCTTACG 1025
DB 972 TCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCCCTCCAAAGAGCTGTGACCTTACG 1031
QY 1026 CAGAAAACAAAGTCTTCACTCGAGTGTGATGATGCTTCTTCTTAATGAGCCCTTA 1085
DB 1032 CAGAAAACAAAGTCTTCACTCGAGTGTGATGATGCTTCTTCTTAATGAGCCCTTA 1091
QY 1086 GATCATTTGTCTCGAAAGCATTTTAAAGCAATTCCTTCATCATGGAATCAATTAACAC 1145
DB 1092 GATCATTTGTCTCGAAAGCATTTTAAAGCAATTCCTTCATCATGGAATCAATTAACAC 1151
QY 1146 GAGTGTGGCTTCCGCTGCTGCTGATGAAGAGGCTCTGAGATCCGACGTGCTCCAAAG 1205
DB 1152 GAGTGTGGCTTCCGCTGCTGCTGATGAAGAGGCTCTGAGATCCGACGTGCTCCAAAG 1211
QY 1206 TCCCTTCCCTCCATCTGATTAACAAACATCTGACATCCCGCTGATATTGGACCTT 1265
DB 1212 TCCCTTCCCTCCATCTGATTAACAAACATCTGACATCCCGCTGATATTGGACCTT 1271
QY 1266 GTGCTTAATGATATCTTCCATGACCAAGCACTCCCTGACGTAAATCCGAGACATCTTTCG 1325
DB 1272 GTGCTTAATGATATCTTCCATGACCAAGCACTCCCTGACGTAAATCCGAGACATCTTTCG 1331
QY 1326 GACTTGCTTGGAGATGTGTTCTTTGTGATCCCTGCACTGATCAGCTGATATACAGA 1385
DB 1332 GACTTGCTTGGAGATGTGTTCTTTGTGATCCCTGCACTGATCAGCTGATATACAGA 1391
QY 1386 GATGCTGTGACACTGTCTATCTTATGAGTTTCGGCACCGGCTCAGTCTTTGAAGAC 1445
DB 1392 ----- 1391
QY 1446 ACGAAGCCGCTTTTGTCAAGCCGACAGCTGATGAAGTCCGCTGTGTGGTGGT 1505
DB 1392 ----- 1391
QY 1506 GCTTCTTAAGAGGGGACATTTGATTTGTTGAAGAGCCACGAGAGAGAGATTAATG 1565
DB 1392 -----GAAGAGCCACGAGAGAGAGATTAATG 1421
QY 1566 AGCCGGAAGATGATGAATATCTGGGCTACTTTTGTCCAAACCGGGAATCTTAATGGGAAC 1625
DB 1422 AGCCGGAAGATGATGAATATCTGGGCTACTTTTGTCCAAACCGGGAATCTTAATGGGAAC 1481
QY 1626 GACTGTCTCTGTGGCCAGCTTATATCTGACTGACAGTATCTTCAGCTGGAATCTTGAAC 1685
DB 1482 GACTGTCTCTGTGGCCAGCTTATATCTGACTGACAGTATCTTCAGCTGGAATCTTGAAC 1541
QY 1686 ATGAGCTCTCGGACAGAGACTCAAAAGAACCGGGGTGGATTTTGTGACAGACCAATGCC 1745
DB 1542 ATGAGCTCTCGGACAGAGACTCAAAAGAACCGGGGTGGATTTTGTGACAGACCAATGCC 1601
QY 1746 CTGATCTGTCTGCTCCGACATGCTCCAGATGCTCTTCTTCTTCTTAATCTTCTCTCT 1805
DB 1602 CTGATCTGTCTGCTCCGACATGCTCCAGATGCTCTTCTTCTTCTTAATCTTCTCTCT 1661
QY 1806 CTCTCCAGCCCTTCTTTCTTTTGTGCTCTTGAAGAGTTATCTTCTGTGATTTTGG 1865

|||||
Db 1662 CTCCTCAGCCTTCTTTCTTTGCTGCTCTGGAAGTATCTTCTGATTTGG 1721
Qy 1866 TTTCCCTCTCTCCCATTAATTTCCCGCATCATCTTACTCTTTGAGCTCAGCTGC 1925
Db 1722 TTTCCCTCTCTCTCCCATTAATTTCTCCGCAATCATCTTCTTTGAGCTCAGCTGC 1781
Qy 1926 TTTCTATGGGAGATCTTGGCAAAACAAGCTGCTTGGCTGATATTTTATGACTTGAAT 1985
Db 1782 TTTCTATGGGAGATCTTGGCAAAACAAGCTGCTTGGCTGATATTTTATGACTTGAAT 1841
Qy 1986 GATCTTACAGAAATTTCTTTCAACATCAAAAGTCAATTTGTCTGAAAGCAA 2040
Db 1842 GATCTTACAGAAATTTCTTTCAACATCAAAAGTCAATTTGTCTGAAAGCAA 1896

RESULT 13
CQ733044 1716 bp DNA linear PAT 03-FRB-2004
LOCUS Sequence 18978 from Patent WO02068579.
DEFINITION CQ733044
ACCESSION CQ733044 GI:42314969
VERSION CQ733044.1 GI:42314969
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 18978 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Source 1. 1716
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 66.0%; Score 1425; DB 6; Length 1716;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 179 AAAAACTGGGCTTCTGCTGAAGGCGCACAGAAACACAGGCTGGGATGATTCAGGG 238
Db 21 AAACAAAGGGCTTCTGCTGAAGGCGCACAGAAACACAGGCTGGGATGATTCAGGG 80
Qy 239 CAAGCAAGTCACTGCTGGGAGGCGCTGCTGCTGGAACGCTTCTCGAGTCCCTT 298
Db 81 CAAGCAAGTCACTGCTGGGAGGCGCTGCTGCTGGAACGCTTCTCGAGTCCCTT 140
Qy 299 TGTGCTCTCCCGCTGGGATCCCTGCAATTTACGAACCGGACGCTGATCCCTGGGA 358
Db 141 TGTGCTCTCCCGCTGGGATCCCTGCAATTTACGAACCGGACGCTGATCCCTGGGA 200
Qy 359 TAACTTGGAGAGCACTCTCTCACTTAATTTGTGCTTCCAGAACTCAGAGTGCCT 418
Db 201 TAACTTGGAGAGCACTCTCTCACTTAATTTGTGCTTCCAGAACTCAGAGTGCCT 260
Qy 419 CTTCAGATCAACAGCTCAAGAGTCAATTAACCGAAATTCGAGGTGCAAGAACTGCT 478
Db 261 CTTCAGATCAACAGCTCAAGAGTCAATTAACCGAAATTCGAGGTGCAAGAACTGCT 320
Qy 479 CTACCTGAACATCTATGGGCTGCGCAAGCGCATCAGGCTCCAGCTCCCGTCTGGT 538
Db 321 CTACCTGAACATCTATGGGCTGCGCAAGCGCATCAGGCTCCAGCTCCCGTCTGGT 380
Qy 539 GTGGTTCCAGAGAGTGCCTTCAAGACTGAGCTCAGCTTCATCTTTGATGGTCCGCCCT 598
Db 381 GTGGTTCCAGAGAGTGCCTTCAAGACTGAGCTCAGCTTCATCTTTGATGGTCCGCCCT 440

Qy 599 GGCTGCTATAGAGAGCTGCTGGTGTGTGCTGCTCAAGTACCGGCTAGGAATAATTTGGTT 658
Db 441 GGCTGCTATAGAGAGCTGCTGGTGTGTGCTGCTCAAGTACCGGCTAGGAATAATTTGGTT 500
Qy 659 CTTCACACATGGGATCAGCATGCTCCGGGAACTGGGCTTTCAAGAGACAGAGTGCCTGC 718
Db 501 CTTCACACATGGGATCAGCATGCTCCGGGAACTGGGCTTTCAAGAGACAGAGTGCCTGC 560
Qy 719 TGTGCTGGGCTGCAAGAAACATGAGTCTTGGTGGGAGCCCGAGCTGTGACCAT 778
Db 561 TGTGCTGGGCTGCAAGAAACATGAGTCTTGGTGGGAGCCCGAGCTGTGACCAT 620
Qy 779 CTTCGGGAGTCCGGGGAGCCCATTAAGTGTCTTCACTTAATCTGTCCCATGAGCCAA 838
Db 621 CTTCGGGAGTCCGGGGAGCCCATTAAGTGTCTTCACTTAATCTGTCCCATGAGCCAA 680
Qy 839 AGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGC 898
Db 681 AGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGC 740
Qy 899 CCATGATTAATGAGAGAGTGAAGAGTCCGAGTGGTTCACATTTCTGTGTAAACAATGC 958
Db 741 CCATGATTAATGAGAGAGTGAAGAGTCCGAGTGGTTCACATTTCTGTGTAAACAATGC 800
Qy 959 GTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTCCAAAGAGCTGCTGAC 1018
Db 801 GTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTCCAAAGAGCTGCTGAC 860
Qy 1019 CCTCAGCCAGAAACAAAGTCTTTCACTGAGGTGATGAGTGTCTTTCTTCTTAATGA 1078
Db 861 CCTCAGCCAGAAACAAAGTCTTTCACTGAGGTGATGAGTGTCTTTCTTCTTAATGA 920
Qy 1079 GCTCTAATATCATATGTCTCAGAAAGCAATTAACCAATTCCTTCATCATCTGAGTCAA 1138
Db 921 GCTCTAATATCATATGTCTCAGAAAGCAATTAACCAATTCCTTCATCATCTGAGTCAA 980
Qy 1139 TAAACAGAGTGTGCTTCTGCTGCTCATGAGAGAGGCTCCTGAGATCTTCAGTGGCTC 1198
Db 981 TAAACAGAGTGTGCTTCTGCTGCTCATGAGAGAGGCTCCTGAGATCTTCAGTGGCTC 1040
Qy 1199 CAACAAATCCCTTGCCTTCATCTGATACAAACATCTCTGACATCCCGCTCAATATT 1258
Db 1041 CAACAAATCCCTTGCCTTCATCTGATACAAACATCTCTGACATCCCGCTCAATATT 1100
Qy 1259 GCACTTGTGGCTAATGAATCTTCATGACAGACATCCCTGAATGAATCCGAGACAG 1318
Db 1101 GCACTTGTGGCTAATGAATCTTCATGACAGACATCCCTGAATGAATCCGAGACAG 1160
Qy 1319 TCTTCTGACCTTCTTGAAGATGTCTTGTGTGCTGCTGACATGATCAAGCTCGATA 1378
Db 1161 TCTTCTGACCTTCTTGAAGATGTCTTGTGTGCTGCTGACATGATCAAGCTCGATA 1220
Qy 1379 TCACAGAGATGTGTGACATCTGTACTTCTATAGATTTGGGACCGGCTCAATGCTT 1438
Db 1221 TCACAGAGATGTGTGACATCTGTACTTCTATAGATTTGGGACCGGCTCAATGCTT 1280
Qy 1439 TGAAGACACGAAGCGGCTTTTGTGAAGCCGACCAAGCTGATGAAGTCCGTTGTGT 1498
Db 1281 TGAAGACACGAAGCGGCTTTTGTGAAGCCGACCAAGCTGATGAAGTCCGTTGTGT 1340
Qy 1499 CGGTGTGCTCTTCTGAAGGGGAGACATTTGTTATGTTCAAGAGGCAAGAGAGAGAGA 1558
Db 1341 CGGTGTGCTCTTCTGAAGGGGAGACATTTGTTATGTTCAAGAGGCAAGAGAGAGAGA 1400
Qy 1559 GTTACTGAGCGGAAAGATGAATACTGGGCTTACCTTTGCTGAAACCGGGA 1611
Db 1401 GTTACTGAGCGGAAAGATGAATACTGGGCTTACCTTTGCTGAAACCGGGA 1453

RESULT 14
AX958392 1857 bp DNA linear PAT 14-JAN-2004
LOCUS AX958392
DEFINITION Sequence 20 from Patent WO0226988.

TITLE Direct Submission
JOURNAL Submitted (04-AUG-2004) Masao Miyazaki, RIKEN, Frontier System
Research, Supra-Biomolecular System Research Group, Hirosewa 2-1,
Wako-shi, Saitama 351-0198, Japan
(E-mail:miyazaki@postman.riken.jp, Tel:01-48-467-9619,
Fax:01-48-467-9620)

FEATURES
Location/Qualifiers
1. 1728

Source
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/chromosome="2"
1. 1728
/gene="cauxin"
1. 1728
/gene="cauxin"
/codon_start=1
/product="carboxylesterase-like urinary excreted protein"
/protein_id="BAD35015.1"
/db_xref="GI:51014275"

/translation="MSGKHGQTLINAVWLAATBEPADAPVRSRLGWKQ
ATVLSSTFVNVFLGIPEAPPLGPRFGRPALMDSRATSPKLCDSVWL
SDQHLKVHPMLVSESDCLYNTIYAPAPANTSKLPVWVPPGAFETGSAIFDGS
ALAAVEDVLIYTOYRLGIFPFTGDOHPNMAFLDQALATVQENIFEGDPH
SVTIFGESGAISSGLVSPMASGLFKHAIMESGVALIPFLAPDEBNDLOYIAR
ICGNVSDSVALLQCLRAKSSRELDINKTKSPTRVYDPPFPDPLDLEKTPNS
IPSVIGVNNHECGFLIPKPEPPEITIGSKSLALHILHRVLIHPQYLYLVADQYFTN
KHSPEIRDSFLDLGDFEVVPGVYTAARHDAAPVTFYFQHPBELKQTRAPV
KADHDEIRFVPGAFILKGNIVMPGATBEELLSRKMRVYANARTDPDREGPLP
MPAYSOBOYRLKLDLNIISVQKLKEQEVFMSDTLPILMSMTAPGPVPILISVL
LPFLPSAP"

ORIGIN

Query Match 55.5%; Score 1198.4; DB 4; Length 1728;
Best Local Similarity 82.7%; Pred. No. 0;
Matches 1370; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 186 GGDCTTCTGCTGAAGGCGCACAGAGGAACACAGGCTGGATTCGATTCAGGCGCAAGCA 245
DB 73 GGTTCAGCTGCTGAATGACCCGGTGAAGACACAGGCTGGATGGTCCGGGGCAAGCA 132
QY 246 GTCACTGCTGGGAAGCCCTGCTGGAAGCGTTCCTGGAAGTCCCTTTGCTGCT 305
DB 133 GCCACTGCTGGGAAGCACTGCTGGAAGCGTTCCTGGAAGTCCCTTTGCTGCT 192
QY 306 CCCCCTGGGAATCCCTGGATTTAGAACCCGCAAGCTGCATCGCTGGATTAATTG 365
DB 193 CCCCCTGGAAGCCCTGGATTTAGAACCAAGCTGCATGCTGGAATGACTCA 252
QY 366 CGAAGAGCACTCTTAACCTTAATTTGCTTCAGAACTCAGAGTGGTCTTAAAT 425
DB 253 CGAGATGCCACCTCTATCTTAATTTGCTTCAGAACTCAGAGTGGTCTTCAAGAT 312
QY 426 CAACATGCTCAAGGTGATTAACCGAATTCGAGGTGCAAGAGACTGCTTAACTG 485
DB 313 CAACACTTCTCAAGGTGATTAACCGAATTTAGAAAGTTCGAAAGACTGCTTAACTT 372
QY 486 AACATCTATGCGCTGCCAGCGCCGATTAAGGCTCCAGCTCCCGCTTGTGTGTTT 545
DB 373 AACATCTATGCGCTGCCAGCGCCGATTAAGGCTCCAGCTCCCGCTTGTGTGTTT 432
QY 546 CAGAGAGGCTCTTCAAGACTGCTCAAGCTCATTTTGAATGGTCCCGCTTGTGCTG 605
DB 433 CTTGAGAGGCTCTTGAAGCTGCTCAAGCTCATTTTGAATGGTCCCGCTTGTGCTG 492
QY 606 TATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 665
DB 493 TATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
QY 666 ACATGGATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
DB 553 ACAGGGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612

QY 726 TGGGTCCAGAAAGATCATGAGTCTTCGCTGGGGAGCCCGAGCTGTGATCACTTTGGC 785
DB 613 TGGGTCCAGAAAGATCATGAGTCTTCGCTGGGGAGCCCGAGCTGTGATCACTTTGGC 672
QY 786 GAGTCCGCGGAGCCATAAGTGTTCAGTCTTAACTGTCTTCCATGGCCAAAGCTTA 845
DB 673 GAGTCCGCGGAGCCATAAGTGTTCAGTCTTAACTGTCTTCCATGGCCAAAGCTTA 732
QY 846 TTCCCAAAAGCCATCATGAGAGTGGGGGCAATCCCTTACTGAGAGCCCATGAT 905
DB 733 TTCCCAAAAGCCATCATGAGAGTGGGGGCAATCCCTTACTGAGAGCCCATGAT 792
QY 906 TATGAGAGAGTGAAGACCTGAGAGTGTGACATTTCTGTGTAAACAATCGTCAGAC 965
DB 793 GATGAGAGAGTGAAGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 852
QY 966 TCTGAGGCTCTGAGAGTGTGACAGCAAAACCTTCAGAGAGCTGTGACCTGAC 1025
DB 853 TCTGAGGCTCTGAGAGTGTGACAGCAAAACCTTCAGAGAGCTGTGACCTGAC 912
QY 1026 CAGAAACCAAACTCTTCACTGAGTGTGATGTGCTTTCTTCTTAATGAGCTCTA 1085
DB 913 AAGAAACCAAACTCTTCACTGAGTGTGATGTGCTTTCTTCTTAATGAGCTCTA 972
QY 1086 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTCATCATGAGAGTCAATAACAC 1145
DB 973 GATCTATTGTCTCAGAAAGCATTTTAAATTAATTCCTCATCATGAGAGTCAATAACAC 1032
QY 1146 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1205
DB 1033 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092
QY 1206 TCCCTTCCCTTCATCTGATTAACAAACATCTGACATTCGCCCTCAGATTTGACCTT 1265
DB 1093 TCCCTTCCCTTCATCTGATTAACAAACATCTGACATTCGCCCTCAGATTTGACCTT 1152
QY 1266 GTGGCTAATGAATCTTCATGACAGACCTCCCTGACTGAATCCGAGACAGTCTTCTG 1325
DB 1153 GTGGCTAATGAATCTTCATGACAGACCTCCCTGACTGAATCCGAGACAGTCTTCTG 1212
QY 1326 GACTTGTGAGATGTGTTCTTGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1385
DB 1213 GACTTGTGAGATGTGTTCTTGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1272
QY 1386 GATGCTGGTGCACCTGTCTACTTCTTAAGATTTGGGACACCGGCTCAGAGCTTTGAAGC 1445
DB 1273 GATGCTGGTGCACCTGTCTACTTCTTAAGATTTGGGACACCGGCTCAGAGCTTTGAAGC 1332
QY 1446 ACGAAGCCGGCTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGCT 1505
DB 1333 ACGAAGCCGGCTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGCT 1392
QY 1506 GCTTCTGAGAGGGGACATTTGTATGTTCGAAGAGCCACGAGAGAGAGAAATTCG 1565
DB 1393 GCTTCTGAGAGGGGACATTTGTATGTTCGAAGAGCCACGAGAGAGAGAAATTCG 1452
QY 1566 AGCGGAAGATGATGAATATCTGCGCTACCTTTGCTCGAAGCGGAAATCTTAATGGAAC 1625
DB 1453 AGCGGAAGATGATGAATATCTGCGCTACCTTTGCTCGAAGCGGAAATCTTAATGGAAC 1512
QY 1626 GACCTGTCTGTGGCAGGCTTATATCTGACTGAGAGTACCTCAGCTGAGCTTGAAC 1685
DB 1513 GACCTGTCTGTGGCAGGCTTATATCTGACTGAGAGTACCTCAGCTGAGCTTGAAC 1572
QY 1686 ATGAGCTCTGAGCAGAGCTCAAGAAACCGCGGTGGAATTTTGAACGAGCACTATCCC 1745
DB 1573 ATGAGCTCTGAGCAGAGCTCAAGAAACCGCGGTGGAATTTTGAACGAGCACTATCCC 1632
QY 1746 CTGATCTGTCTGCTCCGACATGCTCAAGATCTCTTTCTTCTTAATCTTCTCTCT 1805
DB 1633 CTGATCTGTCTGCTCCGACATGCTCAAGATCTCTTTCTTCTTAATCTTCTCTCT 1692
QY 1806 CTCTCTCAGGCTTTCTTTTCTTTTGTGCTCTTGA 1841

Db 1693 GTGCTCTGCGCTTCTTCTATCTCTGCTCTGCTCTGA 1728

Search completed: June 14, 2005, 22:52:42
Job time : 9387.39 sec8

PT surrogate markers, in tissue typing and chromosome mapping.
XX
XX Claim 1; Page 109-110; 122pp; English.

CC The present nucleotide sequence, the coding region of which is also
CC claimed, is that of cDNA encoding human 53010 a novel member of the
CC carboxylesterase family. The invention provides 53010 nucleic acids,
CC antisense molecules, expression vectors, host cells, transgenic animals,
CC 53010 proteins, fusion proteins, antigenic peptides, anti-53010
CC antibodies and methods for detecting the presence of 53010 polypeptides
CC or nucleic acids, of identifying a compound that binds to the 53010
CC polypeptide, and of modulating the activity of the polypeptide. The 53010
CC nucleic acids and polypeptides can act as novel diagnostic and
CC therapeutic agents for controlling disorders involving aberrant or
CC deficient hydrolysis of carboxylic esters. As 53010 mRNA is highly
CC expressed in the central and peripheral nervous system, and its
CC expression is regulated in some rodent pain models, 53010 molecules can
CC also act as novel diagnostic targets and therapeutic agents for
CC controlling neurological disorders, such as pain-related disorders. A
CC claimed method of treating or preventing a disorder (especially a pain-
CC related disorder) characterized by aberrant activity of a 53010-
CC expressing cell involves administering a compound that modulates 53010
CC activity or expression. 53010 nucleic acids are also useful in chromosome
CC mapping, tissue typing, in forensic biology, prognostic assays, in
CC arrays, for detection of variations or mutations, as surrogate markers
CC and in pharmacogenomics
XX
SQ Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;

Query Match 100.0%; Score 2158; DB 6; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGCGCTCGAAGAAACAGGCTGGAGCAATGGAGTAACCAATGTAATTAACCA 60
DB 1 CCAAGCGCTCGAAGAAACAGGCTGGAGCAATGGAGTAACCAATGTAATTAACCA 60
QY 61 TTTAAATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
DB 61 TTTAAATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
QY 121 CTTCAAAATGATGCTTTTCTGATATTCAGAGCCCTGTTGGACAAGAGTGGGAA 180
DB 121 CTTCAAAATGATGCTTTTCTGATATTCAGAGCCCTGTTGGACAAGAGTGGGAA 180
QY 181 AAATGAGGCTTCTGCTGAAGGCGCAAGAGAAACAGAGCTGGAGTGAATTCAGAGCA 240
DB 181 AAATGAGGCTTCTGCTGAAGGCGCAAGAGAAACAGAGCTGGAGTGAATTCAGAGCA 240
QY 241 AACAGTCACTGCTGCTGAAGGCGCAAGAGAAACAGAGCTGGAGTGAATTCAGAGCA 300
DB 241 AACAGTCACTGCTGCTGAAGGCGCAAGAGAAACAGAGCTGGAGTGAATTCAGAGCA 300
QY 301 CTGCTCCCGCGTGGGATCCCTGCGATTTACGAACCCGAGCCTGATGCGCTGGATA 360
DB 301 CTGCTCCCGCGTGGGATCCCTGCGATTTACGAACCCGAGCCTGATGCGCTGGATA 360
QY 361 ACTTGCGAAGACCACTCTCTACCTTAATTTGTGCTCAGAACTCAGAGTGGCTGCT 420
DB 361 ACTTGCGAAGACCACTCTCTACCTTAATTTGTGCTCAGAACTCAGAGTGGCTGCT 420
QY 421 TAGATCAACACATGCTCAAGGTCATTAACCGAAATTTGGAGTGTCAAGAACTGCTCT 480
DB 421 TAGATCAACACATGCTCAAGGTCATTAACCGAAATTTGGAGTGTCAAGAACTGCTCT 480
QY 481 ACCTGAACATCTATGCGCTGCGCAAGCGGATACAGGCTCCAGCTCCCGCTTGGGT 540
DB 481 ACCTGAACATCTATGCGCTGCGCAAGCGGATACAGGCTCCAGCTCCCGCTTGGGT 540
QY 541 GGTTCGCCAGAGGTCCTTCAAGACTGGCTCAGCTCCATCTTTGATGGGTCCGCCCTGG 600
DB 541 GGTTCGCCAGAGGTCCTTCAAGACTGGCTCAGCTCCATCTTTGATGGGTCCGCCCTGG 600

QY 601 CTGCTATAGAGACGTGCTGTTGATGCTCAAGTACCGGCTAGGAATATTGGTTTCT 660
DB 601 CTGCTATAGAGACGTGCTGTTGATGCTCAAGTACCGGCTAGGAATATTGGTTTCT 660
QY 661 TCACCAATGAGATGATGATGCTCCGGGAACTGGGCTTCAAGAACAGAGTGGCTGCTC 720
DB 661 TCACCAATGAGATGATGATGCTCCGGGAACTGGGCTTCAAGAACAGAGTGGCTGCTC 720
QY 721 TGTCTGGGTCAGAAAGATCATGAGTTCTTGGTGGGGAACCCAGCTCTGTGACATCT 780
DB 721 TGTCTGGGTCAGAAAGATCATGAGTTCTTGGTGGGGAACCCAGCTCTGTGACATCT 780
QY 781 TTGGCGAGTCCGCGGAGCCATAGTGTCTTACGTTTACTGTCCTCCATGGCCAAG 840
DB 781 TTGGCGAGTCCGCGGAGCCATAGTGTCTTACGTTTACTGTCCTCCATGGCCAAG 840
QY 841 GCTTATTCACAAACCCATCATGAGAGTGGGTGGCCATCATCTTACCTGAGAGGCC 900
DB 841 GCTTATTCACAAACCCATCATGAGAGTGGGTGGCCATCATCTTACCTGAGAGGCC 900
QY 901 ATGATTAAGAGAGATGAGACCTGACAGTGGTTGACATTTGCTGTGTAAATGGCT 960
DB 901 ATGATTAAGAGAGATGAGACCTGACAGTGGTTGACATTTGCTGTGTAAATGGCT 960
QY 961 CAGACTTGAAGGCTCTGCTGAGTGTCTGAGACAAACCTCCAGAGAGCTGTGACCC 1020
DB 961 CAGACTTGAAGGCTCTGCTGAGTGTCTGAGACAAACCTCCAGAGAGCTGTGACCC 1020
QY 1021 TCAGCCAGAAAAACAAAGCTTTCACTGAGTGTGATGAGGCTTTCTTCTTAATGAGC 1080
DB 1021 TCAGCCAGAAAAACAAAGCTTTCACTGAGTGTGATGAGGCTTTCTTCTTAATGAGC 1080
QY 1081 CTCTAGATCTATGTTCTTCAAGAAACATTTAAAGCAATTCCTTCATCATGAGATCATA 1140
DB 1081 CTCTAGATCTATGTTCTTCAAGAAACATTTAAAGCAATTCCTTCATCATGAGATCATA 1140
QY 1141 AACAGAGTGTGCTTCTGCTGCTAATGAAGAGGCTCTGAGATCTCATGAGTGTGCA 1200
DB 1141 AACAGAGTGTGCTTCTGCTGCTAATGAAGAGGCTCTGAGATCTCATGAGTGTGCA 1200
QY 1201 ACAAGTCCCTTGCCTCCATCTGATTAACAAACATCTGCAATCCGCGCTCAGATTTGCG 1260
DB 1201 ACAAGTCCCTTGCCTCCATCTGATTAACAAACATCTGCAATCCGCGCTCAGATTTGCG 1260
QY 1261 ACCTTGTGCTTAATGATTAATCTTCATGACAGCACTCTGACTGAATTCGAGACAGTC 1320
DB 1261 ACCTTGTGCTTAATGATTAATCTTCATGACAGCACTCTGACTGAATTCGAGACAGTC 1320
QY 1321 TTCTGACTTGTGAGATGTGTCTTGTGGTCCCTGACATGATCAAGCTCGATATTC 1380
DB 1321 TTCTGACTTGTGAGATGTGTCTTGTGGTCCCTGACATGATCAAGCTCGATATTC 1380
QY 1381 ACAGAGTGTGCTGATCACTGTCTTACTTCTTAAGTTTGGGCAACGAGCTCAGTGTCTTG 1440
DB 1381 ACAGAGTGTGCTGATCACTGTCTTACTTCTTGAATTTGGGCAACGAGCTCAGTGTCTTG 1440
QY 1441 AAGACAGAAACCGGCTTTTGTCAAGCGGACCAACCTGATGATGAGTCCGCTTTGTGG 1500
DB 1441 AAGACAGAAACCGGCTTTTGTCAAGCGGACCAACCTGATGATGAGTCCGCTTTGTGG 1500
QY 1501 GTGGTGTCTTCTGAAGGGGGAATTTATGTTTCAGAGGAGCAAGAGAGAGAGAT 1560
DB 1501 GTGGTGTCTTCTGAAGGGGGAATTTATGTTTCAGAGGAGCAAGAGAGAGAGAT 1560
QY 1561 TACTGAGCGGGAAGATGATTAATCTGGGCTTACTTGTCTGAAACCGGAAATCTTAATG 1620
DB 1561 TACTGAGCGGGAAGATGATTAATCTGGGCTTACTTGTCTGAAACCGGAAATCTTAATG 1620
QY 1621 GGAACGACTGTCTCTGTGGCCAGCTTATTAATCTGATGAGCAATCTCCAGCTGGAGT 1680
DB 1621 GGAACGACTGTCTCTGTGGCCAGCTTATTAATCTGATGAGCAATCTCCAGCTGGAGT 1680
QY 1681 TGAACATGAGCTCGGACAGAGATCTCAAGAAACCGCGGTGATTTTGTGACACAGACCA 1740

Db	1681	TGAACATGAGCTCGGACAGAGACTCAAGAACCGCGGGATGATTTTGGACAGACACA	1740A
Qy	1741	TCGCCCTGATCCTGTCTGCGCTCCGACATGCTCCACAGTCTCTTCTTCTTAACTTTC	1800A
Db	1741	TCGCCCGATCCTGTCTGCGCTCCGACATGCTCCACAGTCTCTTCTTCTTAACTTTC	1800A
Qy	1801	TCTCTCTCTCCAGGCTTCTTTTCTTTTGCTCTTGAGAGTTATCTTTCGTGAT	1866A
Db	1801	TCTCTCTCTCCAGGCTTCTTTTCTTTTGCTCTTGAGAGTTATCTTTCGTGAT	1866A
Qy	1861	TTTGCTTCCCTGCTCCGCCATAATCTCCCGCAATCAATAGCTTCTTCTGAGCTCA	1920A
Db	1861	TTTGCTTCCCTGCTCCGCCATAATCTCCCGCAATCAATAGCTTCTTCTGAGCTCA	1920A
Qy	1921	GCTGCTTCTCAATGGGAGCTCTTGCAAAACAAGCTGCTTCCGCTGATATTTATGACCTTA	1980A
Db	1921	GCTGCTTCTCAATGGGAGCTCTTGCAAAACAAGCTGCTTCCGCTGATATTTATGACCTTA	1980A
Qy	1981	GGAATGATCTTACGAAATCTTTTCAACATCAAAAAGTGCATTTGTCCTTGGAAAGCAA	2040A
Db	1981	GGAATGATCTTACGAAATCTTTTCAACATCAAAAAGTGCATTTGTCCTTGGAAAGCAA	2040A
Qy	2041	CAAGATTTCTTCATTAATTTTGGAAAGGCGCTGCTTTATGTTGTCATAATATGTTT	2100A
Db	2041	CAAGATTTCTTCATTAATTTTGGAAAGGCGCTGCTTTATGTTGTCATAATATGTTT	2100A
Qy	2101	TGTAACTCATGTGAAATTAATCAGAATGTAAATAGGAAAAAAAAAAAAAAAAAAAAA	2158A
Db	2101	TGTAACTCATGTGAAATTAATCAGAATGTAAATAGGAAAAAAAAAAAAAAAAAAAAA	2158A

RESULT 2	
ADQ89093	
ID	ADQ89093 standard; cDNA; 2158 BP.
XX	
AC	ADQ89093;
XX	
D7	21-OCT-2004 (first entry)
XX	
DE	Human urological disorder related protein 53010 encoding cDNA SEQ:45
XX	
KM	urological disorder; uropathic; cytostatic; urinary incontinence;
XX	benign prostatic hyperplasia; human; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
PT	96..1841
XX	/*tag= a
XX	/product= "urological disorder related protein 53010"
XX	
PN	WC2004065576-A2.
XX	
PD	
XX	
05	AUG-2004.
XX	
PF	14-JAN-2004; 2004MO-US000750.
XX	
PR	15-JAN-2003; 2003US-0440318P.
XX	
PR	04-FEB-2003; 2003US-044783P.
XX	
PR	27-MAR-2003; 2003US-0457901P.
XX	
PR	08-MAY-2003; 2003US-0468775P.
XX	
PR	19-MAY-2003; 2003US-0471614P.
XX	
PR	16-JUN-2003; 2003US-0478742P.
XX	
PR	18-JUL-2003; 2003US-0488529P.
XX	
PR	30-JUL-2003; 2003US-0491156P.
XX	
PR	02-SEP-2003; 2003US-049594P.
XX	
PR	26-SEP-2003; 2003US-0506332P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
XX	
P1	Karicheti V, Silos-Santiago I, Eliasof SD;
XX	

DR MPI: 2004-562167/54.
 DR P-PSDB: ADQ89094.
 XX Use of polypeptides related to urological disorders, e.g. 44390, 54181,
 PT 211 or for identifying a compound capable of treating a urological
 PT disorder or identifying and treating a subject having a urological
 PT disorder.
 XX
 XX Claim 1; SEQ ID NO 45; 542bp; English.
 CC The present invention describes the use of polypeptides related to
 CC urological disorders for identifying a compound capable of treating a
 CC urological disorder, identifying a subject having a urological disorder,
 CC or treating a subject having a urological disorder. Also described: (1) a
 CC method for identifying a compound capable of treating a urological
 CC disorder; (2) a method for identifying a subject having a urological
 CC disorder; and (3) a method for treating a subject having a urological
 CC disorder. The compound has uropathic and cytostatic activities. The
 CC polypeptides related to urological disorders are useful for identifying a
 CC compound capable of treating a urological disorder, identifying a subject
 CC having a urological disorder, or treating a subject having a urological
 CC disorder. Disorders include urinary incontinence and benign prostatic
 CC hyperplasia. The present sequence encodes a human urological disorder
 CC related protein, which is used in the exemplification of the present
 CC invention.
 XX
 XX Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;
 SQ

Query Match	100:0%	Score 2158;	DB 13;	Length 2158;
Best Local Similarity	100:0%	Prod. No. 0;		
Matches 2158;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	CCAGCGCTCCGAAAAACAGGCTGTGAGAGCATGTGGAGTAAAGCAATATTAATAAAGA	60
Dd	1	CCAGCGCTCCGAAAAACAGGCTGTGAGAGCATGTGGAGTAAAGCAATATTAATAAAGA	60
OY	61	TTTAAAAATTATTTCTTAATAAAAGTAGCAATCCCAATGCCACAGGAACTTATCATCTG	120
Dd	61	TTTAAAAATTATTTCTTAATAAAAGTAGCAATCCCAATGCCACAGGAACTTATCATCTG	120
OY	121	CTTCAACAATGGTCTTTTCTTCTGATTTCTCCAGCCCCCTGTGGGACACAGACATGGGGAA	180
Dd	121	CTTCAACAATGGTCTTTTCTTCTGATTTCTCCAGCCCCCTGTGGGACACAGACATGGGGAA	180
OY	181	AAACTGGGCCCTTCGTGGAAGGGCACAGAGGAACACAAGGCTGGAGATGTATTCAGGGCA	240
Dd	181	AAACTGGGCCCTTCGTGGAAGGGCACAGAGGAACACAAGGCTGGAGATGTATTCAGGGCA	240
OY	241	AGCAAGTCACTGTGCTGTGGAAAGCCCTGTGCTGTGAACGTGTCTCTGAGATCCCTTTG	300
Dd	241	AGCAAGTCACTGTGCTGTGGAAAGCCCTGTGCTGTGAACGTGTCTCTGAGATCCCTTTG	300
OY	301	CTGCTCCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGGATA	360
Dd	301	CTGCTCCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGGATA	360
OY	361	ACTTGCAGAAAGCCACCTCCACCTTAATTTGTGTCCTCCAGAACTCAGAGTGGCTGCT	420
Dd	361	ACTTGCAGAAAGCCACCTCCACCTTAATTTGTGTCCTCCAGAACTCAGAGTGGCTGCT	420
OY	421	TAGATCAACAATGCTCAAGGTGATTAACCGGAATTCGAGTGTCAAGAACTGCTCT	480
Dd	421	TAGATCAACAATGCTCAAGGTGATTAACCGGAATTCGAGTGTCAAGAACTGCTCT	480
OY	481	ACCTGAACATCTATGAGCTGTGCCACGACCGAATACAGGCTCCAGCTCCCGCTTGGTGT	540
Dd	481	ACCTGAACATCTATGAGCTGTGCCACGACCGAATACAGGCTCCAGCTCCCGCTTGGTGT	540
OY	541	GGTTCCAGAGAGTGTCTTAAGCTGATTAACCGGAATTCGAGTGTCAAGAACTGCTCT	600
Dd	541	GGTTCCAGAGAGTGTCTTAAGCTGATTAACCGGAATTCGAGTGTCAAGAACTGCTCT	600
OY	601	CTGCTTATGAGAGTGTGTGTGTGTGTCGTCAGTACCGGCTGAGAAATATTTGGTTCT	660

[illegible]

Db	1681	TGAACATGAGCCTCGGACAGAGACTCAAAGAACCGCGGGGTGAAATTTTGGACCGACGACA	1740
Qy	1741	TCGCCCTGATCTGTCTGTGCTCGCATGCTCCACAGTCCCTTTCTTCCATTAATTCC	1800
Db	1741	TCGCCCTGATCTGTCTGTGCTCGCATGCTCCACAGTCCCTTTCTTCCATTAATTCC	1800
Qy	1801	TCCTCTCTCTCAGGCTTTCTTTTCTTTTGTGCTCTTGAAGATTATCTTCTGTAT	1860
Db	1801	TCCTCTCTCTCAGGCTTTCTTTTCTTTTGTGCTCTTGAAGATTATCTTCTGTAT	1860
Qy	1861	TTTGGTTTCCCTTCTCTCCCAATATTTCTCCGCAATCAATTAAGCTCTTTGAGCTCA	1920
Db	1861	TTTGGTTTCCCTTCTCTCCCAATATTTCTCCGCAATCAATTAAGCTCTTTGAGCTCA	1920
Qy	1921	GCTGCTTTCTATGAGGATCTTGGCAAAACAAGCTGCTTGGCTATTTTATGACTTA	1980
Db	1921	GCTGCTTTCTATGAGGATCTTGGCAAAACAAGCTGCTTGGCTATTTTATGACTTA	1980
Qy	1981	GGAATGATCCCTTAAGAAATTTCTTCAACATCAAAAAGTCATTTGTCTTGAAGGCAA	2040
Db	1981	GGAATGATCCCTTAAGAAATTTCTTCAACATCAAAAAGTCATTTGTCTTGAAGGCAA	2040
Qy	2041	CAGAATTTCTCAATTAATTTGGAAGAGGCTGGCTATTAAGTTGTCATTAATATGTTT	2100
Db	2041	CAGAATTTCTCAATTAATTTGGAAGAGGCTGGCTATTAAGTTGTCATTAATATGTTT	2100
Qy	2101	TGTAACTCATATGAAATTAATCGAATGTAAATATGAAAAAATTTTTTTTTTTTTTTT	2158
Db	2101	TGTAACTCATATGAAATTAATCGAATGTAAATATGAAAAAATTTTTTTTTTTTTTTT	2158
RESULT 3			
AAD40574			
ID	AAD40574	standard; cDNA; 2232 BP.	
XX	AAD40574;		
AC			
XX			
DT	30-OCT-2002	(first entry)	
XX			
DE		Human drug metabolising enzyme (DME-10) cDNA.	
XX			
KM		Human; drug metabolising enzyme; autoimmune; inflammatory disorder;	
KM		acquired immunodeficiency syndrome; AIDS; atherosclerosis; poriasis;	
KM		proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;	
KM		asthma; neurological disorder; Alzheimer's disease; Huntington's disease;	
KM		dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;	
KM		drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;	
KM		renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;	
KM		anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;	
KM		gout; gastrointestinal disorder; gene therapy; virucide; anticoagulant;	
XX		anticonvulsant; nootropic; enzyme; DME-10; gene; ss.	
OS		Homo sapiens.	
XX			
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..1929	
FT		/*tag= a	
FT		/product= "Human DME-10"	
FT	sig_peptide	1..150	
FT		/*tag= b	
FT	mat_peptide	151..1926	
FT		/*tag= c	
FT		/product= "Mature human DME-10"	
XX			
PN	WO200246426-A2.		
PD	13-JUN-2002.		
XX			
PF	04-DEC-2001; 2001WO-US047429.		
XX			
PR	08-DEC-2000; 2000US-0254308P.		
PR	15-DEC-2000; 2000US-0256189P.		
PR	21-DEC-2000; 2000US-0257713P.		

Db	1704	IPATGGGAACGACCTGTTTCTGTGGCAGCCTTAATACTGACTGAGCACTACCTCAAGT	1763
QY	1676	GGACTTTGAACATGAGCCTTCGGACAGAGACTCAAGAACCGCGGGTGGATTATTTGGACGAG	1735
Db	1764	GGACTTTGAACATGAGCCTTCGGACAGAGACTCAAGAACCGGGGGTGGAGTTTGGACGAG	1822
QY	1736	CACCAATCCCCCGATGCTGTGCGCTCCGACATGCTCCACAGTCTCTTTTCTTCTTAAC	1795
Db	1824	CACCAATCCCCCGATGCTGTGCGCTCCGACATGCTCCACAGTCTCTTTTCTTCTTAAC	1883
QY	1796	TTTCTCTCTCTCTCCGACCCCTTTCTTTTCTTTTGTGCTCCTTGAGAGAATATCTTCT	1855
Db	1884	TTTCTCTCTCTCTCCGACCCCTTTCTTTTCTTTTGTGCTCCTTGAGAGAATATCTTCT	1943
QY	1856	GTGATTTTGGTTCCCTCTCTCTCCGCAATATTTCTCCGCAATCATTAAGCTTCTTCTGA	1915
Db	1944	GTGATTTTGGTTCCCTCTCTCTCCGCAATATTTCTCCGCAATCATTAAGCTTCTTCTGA	2003
QY	1916	GCTCAGCTCTTCTATGAGGGATCTTGCAAAACAGATGCTTTCGCGATATTTTAAG	1975
Db	2004	GCTCAGCTCTTCTATGAGGGATCTTGCAAAACAGATGCTTTCGCGATATTTTAAG	2063
QY	1976	ACTTAGGAATGATCTTTACAGAAATTTCTTTTCAACATCAAAAAGTGCATTTGTCTGGAA	2035
Db	2064	ACTTAGGAATGATCTTTACAGAAATTTCTTTTCAACATCAAAAAGTGCATTTGTCTGGAA	2123
QY	2036	GGCAACAGATTTCTCAATAAATTTGGAAAGGGCTGCGCTATTAGTGTCAATATAT	2095
Db	2124	GGCAACAGATTTCTCAATAAATTTGGAAAGGGCTGCGCTATTAGTGTCAATATAT	2183
QY	2096	GGTTTGTAACTCATATGAATAATAAATCAGATGTAAATATGAGAAAAA	2144
Db	2184	GGTTTGTAACTCATATGAATAATAAATCAGATGTAAATATGAGAAAAA	2232

PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen P;
PI Harchshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Deleane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Perilla CH, Anderson SB, Rieux P, Shen EJ, Wu MC, Stive IL;
PI Lagace RE, Spitz PA, Stewart EA, Wingoove J, Velt UA, Kitron ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JT, Gietzen D;
PI Patry S, Shi X, Suarez CJ;

XX
XX
DR WPI; 2004-329368/30.
DR P-PsDB; ABM84114.
XX

PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 1; Page; 190pp; English.
XX
XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorder, or
CC infections caused by virus, bacteria, fungi or parasite. The dthp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dthp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX Sequence 2229 BP; 503 A; 585 C; 547 G; 594 T; 0 U; 0 Other:

Query Match	89.8%	Score 1937.4	DB 13	Length 2229
Best Local Similarity	99.4%	Pred. No. 0		
Matches 1955	Conservative 0	Mismatches 11	Indels 1	Gaps 1
QY	176	GGGAAAACTGGGCGCTTCTGCTGAAGGGCCACAGAGGAAACCAAGGCTGGGATGGAATTCA	235	
DB	264	GGTCCAGAGAGGGCTTCTGCTGAAGGGCCACAGAGAACCAAGGCTGGGATGGAATTCA	323	
QY	236	GGGCAAGCAATCACTGTGTGGGAAAGCCGTGTGTAACGTGTTCTCTGGAGTCCC	295	
DB	324	GGGCAAGCAAGTCACTGTGTGGGAAAGCCGTGTGTAACGTGTTCTCTGGAGTCCC	383	
QY	296	CTTGTGCTGCTCCCGCGTGGGATCCGTCGATTTTAAGAACCCGACGCTGCATCGCCG	355	
DB	384	CTTGTGCTGCTCCCGCGTGGGATCCGTCGATTTTAAGAACCCGACGCTGCATCGCCG	443	
QY	356	GGATTAACCTTGGAGAGGCACCTCTTAACCTTAATTTGTGTGCTTCAGAACTCAGAGTGCT	415	
DB	444	GGATTAACCTTGGAGAGGCACCTCTTAACCTTAATTTGTGTGCTTCAGAACTCAGAGTGCT	503	
QY	416	GGCTTAGATGAACAGATGCTCAAGTGCATTAACCCGAAATTGGAGATGTCAGAAACTG	475	
DB	504	GCTCTTAGATGAACATGATGCTCAAGTGCATTAACCCGAAATTGGAGATGTCAGAAACTG	563	
QY	476	CCTTACCTGAACATCTATGCGCGCTGCGCCAGCGCATACAGGCTCAAGCTCCCGCTTT	535	
DB	564	CCTTACCTGAACATCTATGCGCGCTGCGCCAGCGCATACAGGCTCAAGCTCCCGCTTT	623	
QY	536	GGTGTGTTTCCCAAGAGTGCTTTCAAGACTGAGCTCAGCTCCATCTTTGAATGGTCCG	595	
DB	624	GGTGTGTTTCCCAAGAGTGCTTTCAAGACTGAGCTCAGCTCCATCTTTGAATGGTCCG	683	
QY	596	CCTGTGCTCCTATGAGGACGTGTGTGTTGTGCTGCAGTACCCGGCTAGAAATATTTGG	655	
DB	684	CCTGTGCTCCTATGAGGACGTGTGTGTTGTGCTGCAGTACCCGGCTAGAAATATTTGG	743	
QY	656	TTTCTTACCAACATGGGATCAGCATGTCTCCGGGAAACTGGGCTTCAAGAACCAAGTGGC	715	
DB	744	TTTCTTACCAACATGGGATCAGCATGTCTCCGGGAAACTGGGCTTCAAGAACCAAGTGGC	802	
QY	716	TGCTCTGTCTGGGTCAAGAAAGCAATCGAGTCTTCCGTTGGGGAACCCGAGCTGTGAC	775	
DB	803	TGCTCTGTCTGGGTCAAGAAAGCAATCGAGTCTTCCGTTGGGGAACCCGAGCTGTGAC	862	
QY	776	CATCTTTGGGAGTCCGCGGGAGCCATTAAGTGTTCCTTAATCTTACTGTCTCCCATGGC	835	
DB	863	CATCTTTGGGAGTCCGCGGGAGCCATTAAGTGTTCCTTAATCTTACTGTCTCCCATGGC	922	

QY	836	CAAAAGCTTAATTCACAAAGCCATCATGAGAGTGAGGGGCGCATCATCCCTTACTGGA	895
Db	923	CAAAAGCTTAATTCACAAAGCCATCATGAGAGTGAGGGGCGCATCATCCCTTACTGGA	982
QY	896	GGCCCATGTAATYAGAAAGAGTGAGGACCTGCGAGGTGGTGCACATTTCTGTGTAAACA	955
Db	983	GGCCCATGTAATYAGAAAGAGTGAGGACCTGCGAGGTGGTGCACATTTCTGTGTAAACA	1044
QY	956	TGGCGTCAACTCTGAGGCCCTGCGAGGGTGGCTGAGGACAAACCTCCACAGACCTCT	1015
Db	1043	TGGCGTCAACTCTGAGGCCCTGCGAGGGTGGCTGAGGACAAACCTCCACAGACCTCT	1107
QY	1016	GACCCCTGACGACGAAACAAAGTCTTTGACTCGAGTGGTGTGATGGTCTTTCTCTAA	1075
Db	1103	GACCCCTGACGACGAAACAAAGTCTTTGACTCGAGTGGTGTGATGGTCTTTCTCTAA	1166
QY	1076	TGAGGCTTGAATCTAATGTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT	1133
Db	1163	TGAGGCTTGAATCTAATGTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT	1222
QY	1136	CAATTAACACGAGTGTGGCTTCCTGCTGCGCTATGAAAGAGGCTCCTGAGATCTCAGAG	1199
Db	1223	CAATTAACACGAGTGTGGCTTCCTGCTGCGCTATGAAAGAGGCTCCTGAGATCTCAGAG	1288
QY	1196	CTCCAAACAAGTCCCTTGGCTTCATCTGATAAACAATCTTGACAATCCGCGCTCAGTA	1255
Db	1283	CTCCAAACAAGTCCCTTGGCTTCATCTGATAAACAATCTTGACAATCCGCGCTCAGTA	1344
QY	1256	TTTGCACTTGTGTGCTAATGAAATCTTGCATGACAGACATCCCTGACTGAAATCCGAGA	1315
Db	1343	TTTGCACTTGTGTGCTAATGAAATCTTGCATGACAGACATCCCTGACTGAAATCCGAGA	1402
QY	1316	CAGTCTTCTGACCTTGACCTTGAGAGATGTGTTCTTTGTGTGCTCCCTGACATGATCAAGCTG	1375
Db	1403	CAGTCTTCTGACCTTGACCTTGAGAGATGTGTTCTTTGTGTGCTCCCTGACATGATCAAGCTG	1466
QY	1376	ATATCAACAAGATGTGTGTGCACTGTCTAATCTATGATTTGGGCAACCGGCTCAGTG	1435
Db	1463	ATATCAACAAGATGTGTGTGCACTGTCTAATCTATGATTTGGGCAACCGGCTCAGTG	1522
QY	1436	CTTTGAGAAGACGAAAGCCGGGCTTTTGTCAAAAGCCGACACGCTGATGAAAGTCCGCTTGT	1495
Db	1523	CTTTGAGAAGACGAAAGCCGGGCTTTTGTCAAAAGCCGACACGCTGATGAAAGTCCGCTTGT	1582
QY	1496	GTTTCGGTGTGCTCTTCTCTGAAAGGGGACATTTGTTATGTTGCAAGAGGCCACGAGAGAGA	1555
Db	1583	GTTTCGGTGTGCTCTTCTCTGAAAGGGGACATTTGTTATGTTGCAAGAGGCCACGAGAGAGA	1642
QY	1556	GAACTTACTGAGCCCGGAAGATGATGAAATACCTGGGCTACCTTTGTCTGAACCGGGAATCC	1615
Db	1643	GAACTTACTGAGCCCGGAAGATGATGAAATACCTGGGCTACCTTTGTCTGAACCGGGAATCC	1702
QY	1616	TAAATGGAGACACACCTGTCTCTGTGGGCCAGTTTAACTGACTGAGCAGTACCTCCAGCT	1675
Db	1703	TAAATGGAGACACACCTGTCTCTGTGGGCCAGTTTAACTGACTGAGCAGTACCTCCAGCT	1762
QY	1676	GGACTTGAACATGAGGCTCTGGAACAGAGCTCAAAAGAACCGCGGGGTGATTTTGGACAG	1735
Db	1763	GGACTTGAACATGAGGCTCTGGAACAGAGCTCAAAAGAACCGCGGGGTGATTTTGGACAG	1822
QY	1736	CACCATCCCGCTGATCCGTGTGCTGCTCGGACATGTCCAGAGTCTCTTCTCTCTTCTTAC	1795
Db	1823	CACCATCCCGCTGATCCGTGTGCTGCTCGGACATGTCCAGAGTCTCTTCTCTCTTCTTAC	1882
QY	1796	TTTCTCTCTCTCTCCGACGCTTTCTTTTCTTTGTGTGCTCCTTGAGAAAGTTATCTTCT	1855
Db	1883	TTTCTCTCTCTCTCCGACGCTTTCTTTTCTTTGTGTGCTCCTTGAGAAAGTTATCTTCT	1942
QY	1856	GTAATTTTGTGTTTCCCTCTCTCTCCCAATATTTCTCCGCAATCATTAAGCTTTCTTCTGA	1915
Db	1943	GTAATTTTGTGTTTCCCTCTCTCTCCCAATATTTCTCCGCAATCATTAAGCTTTCTTCTGA	2002

Qy	1916	GCTCAGCTGCTTCTTATGGGGATCTTGCAAAAACAGCTCTTGCCTGATATTTATGG	1975
Db	2003	GCTCAGCTGCTTCTTATGGGGATCTTGCAAAAACAGCTCTTGCCTGATATTTATGG	2062
Qy	1976	ACTTAGGAATGATCCTTACAGAAATCTTTTCAACATCAAAAAGTCAATTTGCTTGAAA	2035
Db	2063	ACTTAGGAATGATCCTTACAGAAATCTTTTCAACATCAAAAAGTCAATTTGCTTGAAA	2122
Qy	2036	GGCAACAAGATTTCTTCAATTAATTTTGGAGAAGGCGTGGCTATTAGTTGTCATTAAT	2095
Db	2123	GGCAACAAGATTTCTTCAATTAATTTTGGAGAAGGCGTGGCTATTAGTTGTCATTAAT	2182
Qy	2096	GCTTTTGTAACATCATATGAATAATTAATCAGAAATGTAATAATAGAAAAA	2142
Db	2183	GCTTTTGTAACATCATATGAATAATTAATCAGAAATGTAATAATAGGATATA	2229
RESULT 5			
ID	ADB62095		
XX	ADB62095	standard; cDNA; 2092 BP.	
XX	ADB62095;		
XX	04-DEC-2003	(first entry)	
DB	Human cDNA encoding clone BRAHM20021910.		
XX	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;		
KW	tissue regeneration; cell regeneration; membrane protein;		
KW	signal transduction-related protein; transcription-related protein;		
KW	osteoporosis; neurological disease; cancer; tumour.		
XX			
OS	Homo sapiens.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	394..1803	
FT		/*tag= a	
FT		/product= "Clone BRAHM20021910 protein"	
XX			
XX	EP1308459-A2.		
PN			
PD	07-MAY-2003.		
XX			
XX	28-MAR-2002; 2002EP-00007401.		
XX			
PR	05-NOV-2001; 2001JP-00379298.		
PR	25-JAN-2002; 2002US-00350978.		
XX			
XX	(HELI-) HELIX RES INST.		
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.		
XX			
PI	Ieogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;		
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;		
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;		
XX			
XX	WPI; 2003-450961/43.		
DR	P-PSDB; ADB64065.		
XX			
PT	New polynucleotides and polypeptides, useful for developing a diagnostic		
PT	marker or medicines for regulation of their expression and activity, or		
PT	as targets of gene therapy.		
XX			
PS	Claim 1; Page; 222pp; English.		
CC	The invention discloses a polynucleotide comprising a sequence selected		
CC	from 1970 fully defined nucleotide sequences which encode novel		
CC	polypeptides. Also claimed is a polypeptide encoded by the polynucleotide		
CC	or its partial peptide, an antibody binding to the polypeptide or peptide		
CC	of the polynucleotide, immunologically assaying the polypeptide or		
CC	peptide of the polynucleotide by contacting the polypeptide or peptide		
CC	with the antibody of the encoded protein, and observing the binding		
CC	between the two, a transformant carrying the polynucleotide in an		
CC	expressible manner and an antisense polynucleotide. The oligonucleotide		

is useful as a primer for synthesizing the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a cDNA of the invention. Note: Some of the sequence data for this patent is not represented in the pruned specification, but is based on sequence information supplied by the European Patent Office.

Sequence 2092 BF; 464 A; 552 C; 497 G; 579 T; 0 U; 0 Other;

Query Match 88.8%; Score 1917; DB 10; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;

```
QY 186 GGGCCCTTCTGCTGAAGGGCCACAGAGAACACCGGCTGGAGTATTCAGGCGAAGCAA 245
DB 138 GGGCCCTTCTGCTGAAGGGCCACAGAGAACACCGGCTGGAGTATTCAGGCGAAGCAA 197
QY 246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAAGTGTCTCTGGAGTCCCTTGTGCTGT 305
DB 198 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAAGTGTCTCTGGAGTCCCTTGTGCTGT 257
QY 306 CCCCCCTGGAGATCCCTCGCATTTTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 365
DB 258 CCCCCCTGGAGATCCCTCGCATTTTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 317
QY 366 CGAAGACCACTCTACCTTAATTT-----GTGGCTCCAAACTCAGAGTGGT 415
DB 318 CGAAGACCACTCTACCTTAATTTTGTAAAGACAGGTGCTTCAGAACTCAGAGTGGT 377
QY 416 GCTCTTAGATCAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAAAGCTG 475
DB 378 GCTCTTAGATCAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAAAGCTG 437
QY 476 CCTCTACCTGAACATTAAGCGCTCCGACGCGATACAGGCTCCAGCTCCCGTCTT 535
DB 438 CCTCTACCTGAACATTAAGCGCTCCGACGCGATACAGGCTCCAGCTCCCGTCTT 497
QY 536 GGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGAGCTCCATCTTGTATGGGTCCGC 595
DB 498 GGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGAGCTCCATCTTGTATGGGTCCGC 557
QY 596 CCTGTGTGCTTAAGAGACGTGTGTTGTGCTGCTCAAGTACCGGCTAAGAAATTTTGG 655
DB 558 CCTGTGTGCTTAAGAGACGTGTGTTGTGCTGCTCAAGTACCGGCTAAGAAATTTTGG 617
QY 656 TTTCTTTCACACATGAGGATCAGATGCTCCGGGAACTGGGCTTCAAGAGACAGATGGC 715
DB 618 TTTCTTTCACACATGAGGATCAGATGCTCCGGGAACTGGGCTTCAAGAGACAGATGGC 677
QY 716 TGTCTGTGCTGGGTCCAGAGAACATGAGTCTTCGTGTGGGAGCCCAAGCTCTGTAGC 775
DB 678 TGTCTGTGCTGGGTCCAGAGAACATGAGTCTTCGTGTGGGAGCCCAAGCTCTGTAGC 737
QY 776 CATCTTTGGCGAGTCCGGGAGCCATAAGTCTTCTAGTCTTAATAGTGTCTCCATGGC 835
DB 738 CATCTTTGGCGAGTCCGGGAGCCATAAGTCTTCTAGTCTTAATAGTGTCTCCATGGC 797
QY 836 CAAAGGCTTAATTCACAAAGCCATCATGAGTGGGGTGGCCATCATCCCTTAACCTGGA 895
DB 798 CAAAGGCTTAATTCACAAAGCCATCATGAGTGGGGTGGCCATCATCCCTTAACCTGGA 857
QY 896 GGGCCATGATTATGGAAGAGTGAAGACCTGCAAGTGTGCACTTTCTGTGTAAACA 955
DB 858 GGGCCATGATTATGGAAGAGTGAAGACCTGCAAGTGTGCACTTTCTGTGTAAACA 917
```

```
QY 956 TGTGTGAGCTGTGAGGCGCTTGAGTGCCTGAGGACAAACCCCTCAAGAGACTGCT 1015
DB 918 TGTGTGAGCTGTGAGGCGCTTGAGTGCCTGAGGACAAACCCCTCAAGAGACTGCT 977
QY 1016 GAGCCCGAGCCGAAACAAAGCTCTTCACTGAGATGTGTGATGCTTTCTTCTTAA 1075
DB 978 GAGCCCGAGCCGAAACAAAGCTCTTCACTGAGATGTGTGATGCTTTCTTCTTAA 1037
QY 1076 TGAAGCTTAGATCTAATGTTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCGAGT 1135
DB 1038 TGAAGCTTAGATCTAATGTTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCGAGT 1097
QY 1136 CAATTAACCAAGTGTGCTTCTGCTGCTATGGAAGAGGCTCTGAGATCTCAAGTG 1195
DB 1098 CAATTAACCAAGTGTGCTTCTGCTGCTATGGAAGAGGCTCTGAGATCTCAAGTG 1157
QY 1196 CTCCAACAAAGTCCCTTGCCCTCAATCTGATCAAAACATCTCGACATCCCGGCTCAFTA 1255
DB 1158 CTCCAACAAAGTCCCTTGCCCTCAATCTGATCAAAACATCTCGACATCCCGGCTCAFTA 1217
QY 1256 TTTGACCTTGTGCTAATGAATCTTCATGACAGCACTCCCTGACTGAATCCGAGA 1315
DB 1218 TTTGACCTTGTGCTAATGAATCTTCATGACAGCACTCCCTGACTGAATCCGAGA 1277
QY 1316 CAGTCTTCTGACATTTGCTTGGAGATGTGTTTGTGTGTCCTTCGACTGATCAAGCTCG 1375
DB 1278 CAGTCTTCTGACATTTGCTTGGAGATGTGTTTGTGTGTCCTTCGACTGATCAAGCTCG 1337
QY 1376 ATATCACAGAGATCTGTGTGACCTGTCTAATCTATGATGTTTGGGACCGGCTCTAGT 1435
DB 1338 ATATCACAGAGATCTGTGTGACCTGTCTAATCTATGATGTTTGGGACCGGCTCTAGT 1397
QY 1436 CTTTGAAGACAGAAAGCGGCTTTTGTCAAGCCGACGAGCTGATGAAGTCCGCTTGT 1495
DB 1398 CTTTGAAGACAGAAAGCGGCTTTTGTCAAGCCGACGAGCTGATGAAGTCCGCTTGT 1457
QY 1496 GTTGTGTGTGTCCTCTCTGAAGGGGGAATTTGTTATGTTGGAAGGACCGAGGAGGA 1555
DB 1458 GTTGTGTGTGTCCTCTCTGAAGGGGGAATTTGTTATGTTGGAAGGACCGAGGAGGA 1517
QY 1556 GAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTAACCCTTGTGGAACCGGGAATCC 1615
DB 1518 GAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTAACCCTTGTGGAACCGGGAATCC 1577
QY 1616 TAAATGGGAAGACCTGTCTCTGTGGGCAAGCTTAATCTGAAGAGATACCTTCCAGCT 1675
DB 1578 TAAATGGGAAGACCTGTCTCTGTGGGCAAGCTTAATCTGAAGAGATACCTTCCAGCT 1637
QY 1676 GAACTTGAACATGAGCCCTGGAAGAGACTCAAGAAACCGGGGTGGAATTTTGAACAG 1735
DB 1638 GAACTTGAACATGAGCCCTGGAAGAGACTCAAGAAACCGGGGTGGAATTTTGAACAG 1697
QY 1736 CACCATCCCTCTGATCTGTGTGCTCTCCAGCATGCTCCAGATGCTCTTCTTTTAC 1795
DB 1698 CACCATCCCTCTGATCTGTGTGCTCTCCAGCATGCTCCAGATGCTCTTCTTTTAC 1757
QY 1796 TTTCTCTCTCTCTCCAGGCTTCTTTTCTTTTGTGTCCTCTGGAAGATTAATCTTCT 1855
DB 1758 TTTCTCTCTCTCTCCAGGCTTCTTTTCTTTTGTGTCCTCTGGAAGATTAATCTTCT 1817
QY 1856 GTGATTTTGTGTTCCCTTCTCTCCAGTAAATTTTCTCCGCAATCATTAAGCTTTCTGA 1915
DB 1818 GTGATTTTGTGTTCCCTTCTCTCCAGTAAATTTTCTCCGCAATCATTAAGCTTTCTGA 1877
QY 1916 GCTCAGCTGTCTTCTAATGGGGAATCTTGAAACCAAGCTGCTTTCCTGATATTTATAG 1975
DB 1878 GCTCAGCTGTCTTCTAATGGGGAATCTTGAAACCAAGCTGCTTTCCTGATATTTATAG 1937
QY 1976 ACTTGAAGATGATCTTTAAGATCTTTTCAACATCAAAAGTGAATTTTGTCTTGA 2035
DB 1938 ACTTGAAGATGATCTTTAAGATCTTTTCAACATCAAAAGTGAATTTTGTCTTGA 1997
QY 2036 GGCACAAAGTTTCTTCAATTAATTTGGAAGAGGCTGACTTAATGTTGCTAATTAAT 2095
```


|||||
DB 1998 GGCACACAAATTTCCCAATTAATTTGGAAGGGCTGGCTTATGTTGATATATAT 2057
QY 2096 GGTTCATCATATGAAATTAATTCAGATGT 2130
DB 2058 GGTTCATCATATGAAATTAATTCAGATGT 2092

RESULT 6
AB086169
ID AB086169 standard; DNA; 1728 BP.
XX
AC AB086169;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human gene. SEQ ID 40.

Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antilicer; vitricide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorder; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.

OS Homo sapiens.
XX
XX WO020250105-A1.
XX
XX 27-JUN-2002.
XX
XX 17-DEC-2001; 2001WO-US049232.
XX
XX 19-DEC-2000; 2000US-0256710P.
XX 20-DEC-2000; 2000US-0257048P.
XX 09-JAN-2001; 2001US-0260482P.
XX 30-JAN-2001; 2001US-0264922P.
XX 06-FEB-2001; 2001US-026797P.
XX 19-MAR-2001; 2001US-0276988P.
XX 04-APR-2001; 2001US-0281535P.
XX 08-MAY-2001; 2001US-0289622P.

(SMK) SMITHKLINE BEECHAM CORP.
PA (SMK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
XX
XX Aggarwal P, Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX WPI; 2002-508784/54.
XX P-PSDB; ABP61004.
XX
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
XX or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX Claim 2(a); Page 251; 335pp; English.

The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, nootropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,

CC cardiant, antilicer, vitricide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records AB086130-AB086184 represent novel human CDNA's
CC of the invention
XX
SQ Sequence 1728 BP; 367 A; 483 C; 437 G; 441 T; 0 U; 0 Other;
Query Match 76.6%; Score 1652.8; DB 6; Length 1728;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

186 GGGCTTCTGCTGAAGGGCCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCAA 245
DB 73 GGGCTTCTGCTGAAGGGCCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCAA 132
QY 246 GTGACTGTCTGGGAAGCCCTGTGCTGTGAAGCTGTTCTTCGAGTCCCTTGTGCT 305
DB 133 GTGACTGTCTGGGAAGCCCTGTGCTGTGAAGCTGTTCTTCGAGTCCCTTGTGCT 192
QY 306 CCCCCGTGGGATCCCTGGATTTTACGAACCCGACGCTGCATCCGCTGGATTAATTG 365
DB 193 CCCCCGTGGGATCCCTGGATTTTACGAACCCGACGCTGCATCCGCTGGATTAATTG 252
QY 366 CGAGAAAGCACTCTTACCTTAATTTGCTCCAGAACTCAGAGTGCCTTACAT 425
DB 253 CGAGAAAGCACTCTTACCTTAATTTGCTCCAGAACTCAGAGTGCCTTACAT 312
QY 426 CAACATCTCTCAGAGTGCATTAACCCGAATTTGAGTGTCAAGAGTGCCTTACCTG 485
DB 313 CAACATCTCTCAGAGTGCATTAACCCGAATTTGAGTGTCAAGAGTGCCTTACCTG 372
QY 486 AACATCTATGCGCCGCGCCACGCGCGATACAGGCTCAAGCTCCGCTTGTGTGCTTC 545
DB 373 AACATCTATGCGCCGCGCCACGCGCGATACAGGCTCAAGCTCCGCTTGTGTGCTTC 432
QY 546 CCAGAGAGTGCCTTCAAGACTGGCTCAGGCTCATCTTGTATGGTCCGCTGCTGCTC 605
DB 433 CCAGAGAGTGCCTTCAAGACTGGCTCAGGCTCATCTTGTATGGTCCGCTGCTGCTC 492
QY 606 TATGAGAGCTGCTGTGTGTGCTGCTCAATACCGGCTAAGAAATTTGTTTCTTAC 665
DB 493 TATGAGAGCTGCTGTGTGTGCTGCTCAATACCGGCTAAGAAATTTGTTTCTTAC 552
QY 666 AATGAGATCAGATGCTCCGAGAACTGGGCTTCAAGAGACAGTGGCTCTCTGCTC 725
DB 553 AATGAGATCAGATGCTCCGAGAACTGGGCTTCAAGAGACAGTGGCTCTCTGCTC 612
QY 726 TGGGTCCAGAAAGATCATGAGTTCCTGCTGGAGAACCCAGCTCTGTGACATCTTGGC 785
DB 613 TGGGTCCAGAAAGATCATGAGTTCCTGCTGGAGAACCCAGCTCTGTGACATCTTGGC 672
QY 786 GAGTCCGCGGAGCATTAAGTCTTCTAGCTTATAGTCTTCCATGGCCAAAGCTTA 845
DB 673 GAGTCCGCGGAGCATTAAGTCTTCTAGCTTATAGTCTTCCATGGCCAAAGCTTA 732
QY 846 TTCACAAAGCATATAGAGAGTGGGTGCGCATATCCCTTACTGAGGCCATGAT 905
DB 733 TTCACAAAGCATATAGAGAGTGGGTGCGCATATCCCTTACTGAGGCCATGAT 792
QY 906 TATGAGAAAGATGAGAGCTGAGGTGTTGACATTTCTGTGTAACAAATGCTCAAG 965

```

Db      793 TATGAGAAAGTGGAGCACTGCAAGTGGTTCACATTTCTGTGTAACAAATGCGTCAAC 852
Qy      966 TCTGAGGCGCTGCTGAGGTGCTTGAGGACAAAACCTTCAAGAGCTGTGACCCCTCAGC 1025
Db      853 TCTGAGGCGCTGCTGAGGTGCTTGAGGACAAAACCTTCAAGAGAGCTGTGACCCCTCAGC 912
Qy      1026 CAGAAAACAAGTCTTCACTCCGAGTGTGATGAGTGTCTTCTTCTTAATGAGCCCTTA 1085
Db      913 CAGAAAACAAGTCTTCACTCCGAGTGTGATGAGTGTCTTCTTCTTAATGAGCCCTTA 972
Qy      1086 GATCTATTTCTCAGAAAAGATTAAAGCAATTCCTTCATTCATCGAGTCAATTAACAC 1145
Db      973 GATCTATTTCTCAGAAAAGATTAAAGCAATTCCTTCATTCATCGAGTCAATTAACAC 1032
Qy      1146 GAGTGTGGCTTCTCTGCTGCTTAAGAGAGGAGCTCTGAGATCTCTCAGTGGCTCAACAG 1205
Db      1033 GAGTGTGGCTTCTCTGCTGCTTAAGAGAGGAGCTCTGAGATCTCTCAGTGGCTCAACAG 1092
Qy      1206 TCCCTTGCCCTCCATCTGATACAAACATCTGCAACATCCCGGCTGAGTATTTGCACTT 1265
Db      1093 TCCCTTGCCCTCCATCTGATACAAACATCTGCAACATCCCGGCTGAGTATTTGCACTT 1152
Qy      1266 GTGGCTTAATGAAATTAATCTTCATGACAAAGCACTCCCTGACTGAAATCCGAGACAGTCTTCTG 1325
Db      1153 GTGGCTTAATGAAATTAATCTTCATGACAAAGCACTCCCTGACTGAAATCCGAGACAGTCTTCTG 1212
Qy      1326 GACTTGGCTGAGAGATGTTGTTCTTGTGTCCTTGGCACTGATCACTGATGATTCACAA 1385
Db      1213 GACTTGGCTGAGAGATGTTGTTCTTGTGTCCTTGGCACTGATCACTGATGATTCACAA 1272
Qy      1386 GATGCTGTGTCACCTGTCTACTTCTATGAGTTTCGCGACCGGCTCAGTGTCTTGAAGAC 1445
Db      1273 GATGCTGTGTCACCTGTCTACTTCTATGAGTTTCGCGACCGGCTCAGTGTCTTGAAGAC 1332
Qy      1446 AGGAACCGGCTTTTGTCAAAGCCGACACAGCTGATGAAAGTCCGCTTGTGTCGGTGT 1505
Db      1333 AGGAACCGGCTTTTGTCAAAGCCGACACAGCTGATGAAAGTCCGCTTGTGTCGGTGT 1392
Qy      1506 GCCTTCCCTGAAGGGGACATTTGTTATGTTGGAAGGACCAAGGAGAGAGAAAGTTACTG 1565
Db      1393 GCCTTCCCTGAAGGGGACATTTGTTATGTTGGAAGGACCAAGGAGAGAGAAAGTTACTG 1452
Qy      1566 ACCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAAACCGGGAATCTTAATGGGAAC 1625
Db      1453 ACCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAAACCGGGAATCTTAATGGGAAC 1512
Qy      1626 GACCTGTCTCTGTGGGCACTTAATCTGACTGAGCACTCACTCACTGAGCTTTGAAC 1685
Db      1513 GACCTGTCTCTGTGGGCACTTAATCTGACTGAGCACTCACTCACTGAGCTTTGAAC 1572
Qy      1686 ATGAGCCTTGGACAGAGACTCAAAAGAACCGGAGTGGATTTTGGACCAAGCAATCCCC 1745
Db      1573 ATGAGCCTTGGACAGAGACTCAAAAGAACCGGAGTGGATTTTGGACCAAGCAATCCCC 1632
Qy      1746 CTGATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCTTAACCTTCTCTCT 1805
Db      1633 CTGATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCTTAACCTTCTCTCT 1692
Qy      1806 CTCTCCAGCCTTTCTTCTTCTTGTGTCCTTA 1841
Db      1693 CTCTCCAGCCTTTCTTCTTCTTGTGTCCTTA 1728

```

RESULT 7
AD53168
ID AD53168 standard; cDNA; 1962 BP.

AC AD53168;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human coding sequence, SEQ ID 736.

```

XX      KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nocotropic;
KW      KM Gene therapy; human; secretory protein; membrane proteins; cancer;
KW      KM Inflammatory disease; osteoporosis; neurological disease; gene; ss.
OS      Homo sapiens.
XX      PN EP1293569-A2.
XX      PD 19-MAR-2003.
XX      PF 21-MAR-2002; 2002EP-00006586.
XX      PR 14-SEP-2001; 2001JP-00328381.
XX      PR 24-JAN-2002; 2002US-0350435P.
XX      PA (HELI-) HELIX RES INST.
XX      PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX      PI Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI      PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI      PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y,
XX      DR MPI; 2003-395539/38.
XX      DR P-PSDB; AD54807.
XX      PT New polynucleotides encoding full-length polypeptides, e.g. secretory
XX      PT and/or membrane proteins, useful for developing medicines for diseases in
XX      PT which the gene is involved, or as target molecules for gene therapy.
XX      PS Claim 1; SEQ ID NO 736; 205pp; English.
XX      CC The present invention relates to novel human secretory or membrane
XX      CC proteins (AD54072-AD545710) and their coding sequences (AD542433-
XX      CC AD544071). The coding sequences are useful in the gene therapy of
XX      CC diseases caused by abnormalities of the proteins, e.g. cancer,
XX      CC inflammatory diseases, osteoporosis or neurological disease.
SQ      Sequence 1962 BP; 428 A; 541 C; 473 G; 520 T; 0 U; 0 Other;

Query Match 73.2%; Score 1579.4; DB 10; Length 1962;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

Qy      186 GGGCTTCTGCTGAAGGCCACAGAGAACCAACAGCTGGATGATTCAGGCAAGCAA 245
Db      222 GGGCTTCTGCTGAAGGCCACAGAGAACCAACAGCTGGATGATTCAGGCAAGCAA 281
Qy      246 GTCACTGTGCTGGAAGCCCTGTGCTGTGAACGTTTCTCGGAAGTCCCTTTGTGCTGT 305
Db      282 GTCACTGTGCTGGAAGCCCTGTGCTGTGAACGTTTCTCGGAAGTCCCTTTGTGCTGT 341
Qy      306 CCCCCTGTGGATTCCTTGCATTTTACGAACCCGACGCTGATCGCCCTGGATTAACCTTG 365
Db      342 CCCCCTGTGGATTCCTTGCATTTTACGAACCCGACGCTGATCGCCCTGGATTAACCTTG 401
Qy      366 CGAAGACCACTTCATACCTTAATTTTGTGCTCCAGAACTCAAGTGGCTGCTTAAGT 425
Db      402 CGAAGACCACTTCATACCTTAATTTTGTGCTCCAGAACTCAAGTGGCTGCTTAAGT 461
Qy      426 CAACACATGCTCAAGTGCATTAACCGGAATTCGAGATGCAGAAAGTGCCTCTACCTG 485
Db      462 CAACACATGCTCAAGTGCATTAACCGGAATTCGAGATGCAGAAAGTGCCTCTACCTG 521
Qy      486 AACATCTATGCGCTGCTCCACGCGGATACAGGCTCCAGCTCCCGCTTTGTGTGTTTC 545
Db      522 AACATCTATGCGCTGCTCCACGCGGATACAGGCTCCAGCTCCCGCTTTGTGTGTTTC 581
Qy      546 CCAAGAGTGCCTTTCAAGACTGAGCTCAGGCTCCATCTTTGATGGGTCGGCCCTGGCTGCC 605
Db      582 CCAAGAGTGCCTTTCAAGACTGAGCTCAGGCTCCATCTTTGATGGGTCGGCCCTGGCTGCC 641
Qy      606 TATGAGGACGTGTGTTGTGTGTCAGTACCGGCTGAGAAATATTTGTTCTTCAACC 665

```

```

Db 642 TATGAGAGAGTGTGTTGTTGTCGTCAGATACCGCTAGAAATATTTGTTTCTTCAAC 701
Qy 666 AATAGGATACAGATGCTCCGGGGAATGAGCTTCAAGACCAAGTGGCTGTCTGTGC 725
Db 702 AATAGGATACAGATGCTCCGGGGAATGAGCTTCAAGACCAAGTGGCTGTCTGTGC 761
Qy 726 TGGGTCAGAGAAACATGAGTCTTCCGGTGGGGAACCCAGCTCTGTGACCAATCTTTGGC 785
Db 762 TGGGTCAGAGAAACATGAGTCTTCCGGTGGGGAACCCAGCTCTGTGACCAATCTTTGGC 821
Qy 786 GAGTCGCGGGAAGCCATTAAGTGTTCCTTAATCTGTCTTCAATGAGCCAAAGGCTTA 845
Db 822 GAGTCGCGGGAAGCCATTAAGTGTTCCTTAATCTGTCTTCAATGAGCCAAAGGCTTA 881
Qy 846 TTCACCAAGGCAATCATGAGAGTGGGTGGCCATATCCCTTAAGCTGAGGCGCCATGAT 905
Db 882 TTCACCAAGGCAATCATGAGAGTGGGTGGCCATATCCCTTAAGCTGAGGCGCCATGAT 941
Qy 906 TATGAGAAAGTGAAGACCTGAGAGTGTGTCATTTCTGTGTAAACAATGGCTGAGAC 965
Db 942 TATGAGAAAGTGAAGACCTGAGAGTGTGTCATTTCTGTGTAAACAATGGCTGAGAC 1001
Qy 966 TCTGAGGCGCTGCTGAGAGTGTGTCAGACAAACCTTCAAGAGAGCTGTGACCTCAGC 1025
Db 1002 TCTGAGGCGCTGCTGAGAGTGTGTCAGACAAACCTTCAAGAGAGCTGTGACCTCAGC 1061
Qy 1026 CAGAAAACAAAGCTTTTCACTGAGTGTGTTGATGTCCTTTCTTCTTAATAGCTCTTA 1085
Db 1062 CAGAAAACAAAGCTTTTCACTGAGTGTGTTGATGTCCTTTCTTCTTAATAGCTCTTA 1121
Qy 1086 GATCTATTGTCTCAGAAACATTTAAAGAAATTCCTTCATCATATGGAGTCAATACCAAC 1145
Db 1122 GATCTATTGTCTCAGAAACATTTAAAGAAATTCCTTCATTAATGAGAGTCAATACCAAC 1181
Qy 1146 GAGTGTGCTTCTGCTGCTTCTGATGAGAGAGGCTCTGAGATCTCAGTGGCTTCAACAAG 1205
Db 1182 GAGTGTGCTTCTGCTGCTTCTGATGAGAGAGGCTCTGAGATCTCAGTGGCTTCAACAAG 1241
Qy 1206 TCCCTTGGCTTCCATCTGATTAACAAACATCTGACATCCCGCTCAATATTTGCACTTT 1265
Db 1242 TCCCTTGGCTTCCATCTGATTAACAAACATCTGACATCCCGCTCAATATTTGCACTTT 1301
Qy 1266 GTGGCTAATGATCTTCCATGACAAAGCACTCCCTGACGTGAATCCGAGACAGTCTTGG 1325
Db 1302 GTGGCTAATGATCTTCCATGACAAAGCACTCCCTGACGTGAATCCGAGACAGTCTTGG 1361
Qy 1326 GACTTGTGAGAGATGTTCTTGTGTGTCCCTGCACTGATCAAGCTGATATCAAGAGA 1385
Db 1362 GACTTGTGAGAGATGTTGTGTGTGTCCCTGCACTGATCAAGCTGATATCAAGAGA 1421
Qy 1386 GATGCTGTGCACTGTCTACTTCTAATGAGTTCCGACACGGCTCAGTGTCTTGAAGAC 1445
Db 1422 ----- 1421
Qy 1446 ACGAAGCGGCTTTTGTCAAAAGCCAGACAGCTGATGAAGTCCGCTTGTGTGCTGTGT 1505
Db 1422 ----- 1421
Qy 1506 GCTTCTGGAAGGGGACATGTTGTTATGTTGAAGAGACCAAGGAGAGAGAGTTACTG 1565
Db 1422 -----GAAAGAGCAACGAGAGAGAGAGAGTTACTG 1451
Qy 1566 AGCCGGAAGATGATGAATACTGCGGCTACTTGTCTGAAACCGGGAATCTTAATGGGAAC 1625
Db 1452 AGCCGGAAGATGATGAATACTGCGGCTACTTGTCTGAAACCGGGAATCTTAATGGGAAC 1511
Qy 1626 GACCTGTCTGTGTGCGAGCTTAAATCTGATGAGCAATCACTCAAGTGTGAACCTTGAAC 1685
Db 1512 GACCTGTCTGTGTGCGAGCTTAAATCTGATGAGCAATCACTCAAGTGTGAACCTTGAAC 1571
Qy 1686 ATGAGCTTGAAGAGACTCAAGAAACCGCGGTTGATTTTGTGACAGCAATCCCC 1745

```

```

Db 1572 ATGAGCTTGAAGAGACTCAAGAAACCGCGGTTGATTTTGTGACAGACCAATCCCC 1631
Qy 1746 CTGATCTGTCTGCGCTCCGACAGATGCTCCCAAGTCCCTCTTCTTCTTAATCTTCTCT 1805
Db 1632 CTGATCTGTCTGCGCTCCGACAGATGCTCCCAAGTCCCTCTTCTTCTTAATCTTCTCT 1691
Qy 1806 CTCCTTCAGACCTTCTTCTTCTTGTGTGCTCCCTGAGAAAGTATCTTCTGTGATTTTGG 1865
Db 1692 CTCCTTCAGACCTTCTTCTTCTTGTGTGCTCCCTGAGAAAGTATCTTCTGTGATTTTGG 1751
Qy 1866 TTTCCCTTCTCTCCCAATATTTCTCCGCAATCATTAAGCTTCTTCTGAGCTCAGCTGC 1925
Db 1752 TTTCCCTTCTCTCCCAATATTTCTCCGCAATCATTAAGCTTCTTCTGAGCTCAGCTGC 1811
Qy 1926 TTTCTAATGGGGAATCTTGAAGAAACAAGCTGCTGTGATATTTTAATGACTTAGGAAT 1985
Db 1812 TTTCTAATGGGGAATCTTGAAGAAACAAGCTGCTGTGATATTTTAATGACTTAGGAAT 1871
Qy 1986 GATCTTACAGAAATCTTTCACATCAAAAGTGAATTTGCTTGAAGGCAACAAGA 2045
Db 1872 GATCTTACAGAAATCTTTCACATCAAAAGTGAATTTGCTTGAAGGCAACAAGA 1931
Qy 2046 TTTCTTCAATTAATTTGGAAGAGGCTGACC 2076
Db 1932 TTTCTTCAATTAATTTGGAAGAGGCTGACC 1962

RESULT 8
ABX72267
ID ABX72267 standard; cDNA, 1746 BP.
XX
AC ABX72267;
XX
XX 03-JUN-2003 (first entry)
DE
DE Human NOVX polynucleotide #98.
XX
XX Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;
XX hypertension; congenital heart defect; aortic stenosis; valve disease;
XX atrial septal defect; atriocentricular canal defect; ductus arteriosus;
XX pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
XX tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
XX obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
XX Parkinson's disease; immune disorder; haematopoietic disorder;
XX hemophilia; hypercoagulation; Crohn's disease; cancer.
OS Homo sapiens.
PN WO200281498-A2.
XX
XX 17-OCT-2002.
PD
PD 03-APR-2002; 2002WO-US010780.
XX
XX 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.

```

PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
(CURA-) CURAGEN CORP.

PI Guo X, Kekuda R, Miller CE, Malyanker UM, Spytek KA,
PI Paturajan M, Liu X, Gusev VY, Li L, Vernet CM, Zerhusen BD,
PI German L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V,
PI Padigaru M, Shimkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W,
PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ,
PI McDougall JR, Rothenberg ME, Mazur A, Millet I, Payman JA,
PI Ellerman K;

DR MPI: 2003-046858/04.
DR P-PSDB; ABUS4639.

PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.

XX Claim 17, Page 303; 666p; English.

XX The invention relates to human polypeptides, termed NOVX, and the
XX polynucleotides encoding them. The polypeptides and polynucleotides are
XX useful for diagnosing disease, and screening for potential therapeutic
XX agents. The sequences are useful for treating metabolic disorders,
XX cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
XX stenosis, atrial septal defect (ASD), atrioventricular canal defect,
XX ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
XX septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
XX atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
XX disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
XX haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
XX and cancer. Sequences ABX72170-ABX72275 represent human NOVX
XX polynucleotides of the invention

XX Sequence 1746 BP; 402 A; 457 C; 459 G; 428 T; 0 U; 0 Other;

Query Match 66.5%; Score 1435.6; DB 8; Length 1746;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1196; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

QY 176 GGGAAAAAAGTGGGCTTTGCTGAGAGGCGACAGAGAAACACAGGCTGGAGTGTATCA 235
Db 69 GGTCAAGAGAGGCGCTTGTGCTGAGAGGCGACAGAGAAACACAGGCTGGAGTGTATCA 128
QY 236 GGGCAAGCAAGTCAATGTCGTCGGAAGCCCTGTGCTGTGAACGTGTTCTCGAGTCC 295
Db 129 GGGCAAGCAAGTCAATGTCGTCGGAAGCCCTGTGCTGTGAACGTGTTCTCGAGTCC 188
QY 296 CTTTGCTGCTCCCGCGTGGGATCCCTGCGATTACGAAACCCGACCTCGATGGCCTTG 355
Db 189 CTTTGCTGCTCCCGCGTGGGATCCCTGCGATTACGAAACCCGACCTCGATGGCCTTG 248
QY 356 GGATTAATTGCGAAGGCACTCTCACTTAATTGCTGCTCGAAGTCAAGTGGCT 415
Db 249 GGATTAATTGCGAAGGCACTCTCACTTAATTGCTGCTCGAAGTCAAGTGGCT 308

QY 416 GCTCTTAGATCAACATGCTCAAGTGCATACCGGAATTCGGAGTGTGAGAGACTG 475
Db 309 GCTCTTAGATCAACATGCTCAAGTGCATACCGGAATTCGGAGTGTGAGAGACTG 368
QY 476 CCTCTACCTGAACATCTATGCGCTGCGCCAGCGATACAGGCTCCAAAGTCCCGTCT 535
Db 369 CCTCTACCTGAACATCTATGCGCTGCGCCAGCGATACAGGCTCCAAAGTCCCGTCT 428
QY 536 GGTGTGTGTTCCGAGAGGTCCTTCAAGATGCTCAGCTCCATCTTGTATGGTCCGC 595
Db 429 GGTGTGTGTTCCGAGAGGTCCTTCAAGATGCTCAGCTCCATCTTGTATGGTCCGC 488
QY 596 CCTGTGCTGTATGAGAGATGTCGTGTTGTGCTGCTCCAGTACCGGCTAGAAATTTTG 655
Db 489 CCTGTGCTGTATGAGAGATGTCGTGTTGTGCTGCTCCAGTACCGGCTAGAAATTTTG 548
QY 656 TTTCTTCAACACATGGGATTCAGATGCTCCGGGGAACCTGGGCTTCAAGAACAGGTGCG 715
Db 549 TTTCTTCAACACATGGGATTCAGATGCTCCGGGGAACCTGGGCTTCAAGAACAGGTGCG 608
QY 716 TGCTCTGTCTGGGTCCAGAAAGCAATGAGTTCTTGCGTGGGAGCCAGCTGTGAGC 775
Db 609 TGCTCTGTCTGGGTCCAGAAAGCAATGAGTTCTTGCGTGGGAGCCAGCTGTGAGC 668
QY 776 CATCTTTGGCGAGTCCCGGGAGCCATTAAGTGTCTTAAGTCTTAATCTGTCTCCATGCG 835
Db 669 CATCTTTGGCGAGTCCCGGGAGCCATTAAGTGTCTTAAGTCTTAATCTGTCTCCATGCG 728
QY 836 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATTCATCCCTTACCTGGA 895
Db 729 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATTCATCCCTTACCTGGA 788
QY 896 GAGCCATGATTATGAGAAAGTGAAGACCTGACAGTGTGACATTTCTGTGTTAACAA 955
Db 789 GAGCCATGATTATGAGAAAGTGAAGACCTGACAGTGTGACATTTCTGTGTTAACAA 848
QY 956 TGCGTCAGACTCTGAGGCGCTGCTGAGTGTGCTGAGGACAAACCTTCAGAGAGTGTCT 1015
Db 849 TGCGTCAGACTCTGAGGCGCTGCTGAGTGTGCTGAGGACAAACCTTCAGAGAGTGTCT 908
QY 1016 GACCTCAGGCGAGAAACAAAGTCTTCACTGAGTGTGATGATGCTTCTTCTCTTAA 1075
Db 909 GACCTCAGGCGAGAAACAAAGTCTTCACTGAGTGTGATGATGCTTCTTCTCTTAA 968
QY 1076 TGACCTCTGATCTTATGCTCTGAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT 1135
Db 969 TGACCTCTGATCTTATGCTCTGAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT 1028
QY 1136 CAATTAACACAGAGTGGCTTCCGCTGCTATGAGAGAGGCTCCGAGATCTCTAGTGG 1195
Db 1029 CAATTAACACAGAGTGGCTTCCGCTGCTATGAGAGAGGCTCCGAGATCTCTAGTGG 1088
QY 1196 CTCGAACAGTCCCTTGCCCTCATCTGATACAA-----ACATCTGCA 1240
Db 1089 CACTCCCTCAACACGATGATCAGCTTGGCTTCAACAGCTGGGCAATTCACAGAAAGCA 1148
QY 1241 CATCCGCTCAGTATTTTGAACCTTGTGGTAAATGAAATCTTCATGACAGACATCCCT 1300
Db 1149 TATCCGCTCAGTATTTTGAACCTTGTGGTAAATGAAATCTTCATGACAGACATCCCT 1208
QY 1301 GACTGAATCCGAGAGTCTTCTGAGACTTGTGAGATGATGTTCTTGTGGTCCCTG 1360
Db 1209 GACTGAATCCGAGAGTCTTCTGAGACTTGTGAGATGATGTTCTTGTGGTCCCTG 1268
QY 1361 ACTGATCAAGCTCGATATCAACAGAGATCTGTGACCTGTCTACTTATGAGATTTTG 1420
Db 1269 ACTGATCAAGCTCGATATCAACAGAGATCTGTGACCTGTCTACTTATGAGATTTTG 1328
QY 1421 GCACGGGCTCAGTGTCTTGAAGACAGAGCGGCTTTGTCAAAAGCGACACAGCTGA 1480
Db 1329 GCACGGGCTCAGTGTCTTGAAGACAGAGCGGCTTTGTCAAAAGCGACACAGCTGA 1388

QY 1481 TGAAGTCGGCTTGTGTCGGTGGCTTCTGAAAGGGGAGACTTGTATGTTGAAAG 1540
DB 1389 TGAAGTCGGCTTGTGTCGGTGGCTTCTGAAAGGGGAGACTTGTATGTTGAAAG 1448
QY 1541 AGCCACGAG 1600
DB 1449 AGCCACGAG 1508
QY 1601 TGAACCGGGAATCTTAATGGAACGAGCTGTCTGTGGCCGCTTATATCTGACTGA 1660
DB 1509 TGAACCGGGAATCTTAATGGAACGAGCTGTCTGTGGCCGCTTATATCTGACTGA 1568
QY 1661 GCAATACCTTCAGCTGAGCTTGAACATGAGCTCGAGCAGAGACTCAAGAGACCGCGAGT 1720
DB 1569 GCAATACCTTCAGCTGAGCTTGAACATGAGCTCGAGCAGAGACTCAAGAGACCGCGAG 1628
QY 1721 GGAATTTGG 1730
DB 1629 AGATGTGTGG 1638

RESULT 9
ABQ86170
ID ABQ86170 standard; DNA; 1746 BP.
XX
AC ABQ86170;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human gene. SEQ ID 41.
XX
XX Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;
XX neurotrophic; neuroprotective; immunosuppressive; haemostatic;
XX antiinflammatory; cardiac; antitumor; antiviral; antithyroid;
XX cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
XX wound healing disorders; atherosclerosis; Parkinson's disease;
XX Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
XX inflammation; neoplastic disease; nervous system disorder;
XX cardiovascular disorders; pancreatitis; respiratory disorder;
XX hyperproliferation; systemic autoimmune disease; hyper-immunity;
XX developmental abnormality; gastrointestinal ulceration; neuropathy;
XX haematological abnormality; metabolic disease; sperm dysfunction;
XX thyroid disorder; hypothyroidism; brain damage; colitis;
XX cone photo-transduction deficiency; neurological disease; stroke;
XX anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
XX trachea; thymus; lymph node; muscular system; obesity; anorexia;
XX growth abnormality; precocious puberty; gene; ss.
OS Homo sapiens.
XX
XX MO200250105-A1.
PN
XX
XX 27-JUN-2002.
PD
PF 17-DEC-2001; 2001MO-US049232.
XX
XX 19-DEC-2000; 2000US-0256710P.
XX 20-DEC-2000; 2000US-0257048P.
XX 29-JAN-2001; 2001US-0260482P.
XX 30-JAN-2001; 2001US-0264922P.
XX 06-FEB-2001; 2001US-0266797P.
XX 19-MAR-2001; 2001US-0276988P.
XX 04-APR-2001; 2001US-0281535P.
XX 08-MAY-2001; 2001US-0289622P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAXO) GLAXO GROUP LTD.
XX
XX Agatwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q,
XX WPI; 2002-508784/54.

DR P-PSDB; ABP61005.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
PS Claim 2(a); Page 251-252; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotrophic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiac, anticancer, antiviral, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
XX of the invention
XX
SQ Sequence 1746 BP; 401 A; 464 C; 459 G; 422 T; 0 U; 0 Other;
XX
XX Query Match 66.4%; Score 1432.4; DB 6; Length 1746;
XX Best Local Similarity 95.4%; Pred. No. 0;
XX Matches 1494; Conservative 0; Mismatches 51; Indels 21; Gaps 1;
XX
QY 186 GGGCTTTGCTGAAAGGGCCAGAGAGAACACAGAGCTGGAGTGAATTCAGGGCAAGCAA 245
DB 73 GGGCTTTGCTGAAAGGGCCAGAGAGAACACAGAGCTGGAGTGAATTCAGGGCAAGCAA 132
QY 246 GTCACTGTGCTGGAAGCCCTGTGCTGTGAACGTGTCTCGAGTCCCTTTGCTGCT 305
DB 133 GTCACTGTGCTGGAAGCCCTGTGCTGTGAACGTGTCTCGAGTCCCTTTGCTGCT 192
QY 306 CCCCCTGGGATCCCTGCGATTTACGAACCCGAGCTGTCAGTCCCTGGATTAATTG 365
DB 193 CCCCCTGGGATCCCTGCGATTTACGAACCCGAGCTGTCAGTCCCTGGATTAATTG 252
QY 366 CGAAGAGCCACCTCCATACCTATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGAT 425
DB 253 CGAAGAGCCACCTCCATACCTATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGAT 312
QY 426 CAACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAGACTGCTCTACTG 485
DB 313 CAACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAGACTGCTCTACTG 372
QY 486 AACATCTATGCGCTGCCCAAGCGATTAAGGCTCAAGTCCCGTCTGTGTGGTTG 545
DB 373 AACATCTATGCGCTGCCCAAGCGATTAAGGCTCAAGTCCCGTCTGTGTGGTTG 432
QY 546 CCAAGAGAGCTCTTAAGAGCTGCTCAAGCTCAATTTGATAGGTCCTGGCTGGCTGCC 605
DB 433 CCAAGAGAGCTCTTAAGAGCTGCTCAAGCTCAATTTGATAGGTCCTGGCTGGCTGCC 492
QY 606 TATGAGAGAGTGTGTTGTGCTGCTCAAGTCCGCTAGAAATATTTGGTTCTTAC 665
DB 493 TATGAGAGAGTGTGTTGTGCTGCTCAAGTCCGCTAGAAATATTTGGTTCTTAC 552
QY 666 ACATGGAGATCAGATCTCCGGGAACTGGGCTTCAAGAGCAGAGTGTGCTGTCTC 725
DB 553 ACATGGAGATCAGATCTCCGGGAACTGGGCTTCAAGAGCAGAGTGTGCTGTCTC 612

726 TGGGTCAGAGAAATGAGTCTTCCGTGGGGAGCCAGCTCTGTGACATCTTTGGC 785
 613 TGGGTCAGAGAAATGAGTCTTCCGTGGGGAGCCAGCTCTGTGACATCTTTGGC 672
 786 GAGTCCGGGGAGCCATTAAGTCTTCTAGTCTTAATAGTCTCCCATGGCCAAAGGCTTA 845
 673 GAGTCCGGGGAGCCATTAAGTCTTCTAGTCTTAATAGTCTCCCATGGCCAAAGGCTTA 732
 846 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGGCCCATGAT 905
 733 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGGCCCATGAT 792
 906 TATGAGAAGAGTGAAGACCTGACAGGTGGTTCACATTTCTGTGTAAACAATGCTTCAGAC 965
 793 TATGAGAAGAGTGAAGACCTGACAGGTGGTTCACATTTCTGTGTAAACAATGCTTCAGAC 852
 966 TCTGAGGCCCTGCTGAGGTGCTGAGGACAAACCCCTCAAGAGCTGTGACCCCTCAGC 1025
 853 TCTGAGGCCCTGCTGAGGTGCTGAGGACAAACCCCTCAAGAGCTGTGACCCCTCAGC 912
 1026 CAGAAAACAAAGTCTTCACTCCGAGTGTGATGAGTCTTCTTCTTAATGAGCTCTA 1085
 913 CAGAAAACAAAGTCTTCACTCCGAGTGTGATGAGTCTTCTTCTTAATGAGCTCTA 972
 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACAC 1145
 973 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACAC 1032
 1146 GAGTGGGCTTCTGCTGCTCTTAAGAGAGGCTCTGAGATCTCAGTGGCTCCAAACAG 1205
 1033 GAGTGGGCTTCTGCTGCTCTTAAGAGAGGCTCTGAGATCTCAGTGGCTCCAAACAG 1092
 1206 TCCCTT-----GCCCTCATCTGATACAAACATCCGACATC 1244
 1093 AACCGGATGACAGCTTTGGCTTCAACAGCTGGGCAATTCACAGAAAGATCAGCAATC 1152
 1245 CCGCCTCAGATTGTCACCTTGGGCTTAATGAATCTTCCATGACAGCACTCCCTGACT 1304
 1153 CCGCCTCAGATTGTCACCTTGGGCTTAATGAATCTTCCATGACAGCACTCCCTGACT 1212
 1305 GAAATCCGAGACAGTCTTGTGACCTTGTGAGAGTGTCTTGTGCTTCCCTGACATC 1364
 1213 GAAATCCGAGACAGTCTTGTGACCTTGTGAGAGTGTCTTGTGCTTCCCTGACATC 1272
 1365 ATCAGAGCTGATATCAACAGAGTGTGTCACCTGTCTACTTCTATGAGTTCCGACAC 1424
 1273 ATCAGAGCTGATATCAACAGAGTGTGTCACCTGTCTACTTCTATGAGTTCCGACAC 1332
 1425 CCGCCTCAGATTGTCACCTTGTGAGAGTGTCTTGTGCTTCCCTGACATC 1484
 1333 CCGCCTCAGATTGTCACCTTGTGAGAGTGTCTTGTGCTTCCCTGACATC 1392
 1485 GTCGCGTTTGTGTTCCGTGTGCTCTGAAAGGGGGAATTGTTATGTTGAGAGAGCC 1544
 1393 GTCGCGTTTGTGTTCCGTGTGCTCTGAAAGGGGGAATTGTTATGTTGAGAGAGCC 1452
 1545 ACGAGAGAGAGAAATTAATGAGCCGGAAGATGATGAATTAATCTGAGAGAGAG 1604
 1453 ACGAGAGAGAGAAATTAATGAGCCGGAAGATGATGAATTAATCTGAGAGAGAG 1512
 1605 ACGGGAATCTTAATGAGAGACAGCTGTCTGTGGCCAGCTTAATTAATCTGAGAGAG 1664
 1513 ACGGGAATCTTAATGAGAGACAGCTGTCTGTGGCCAGCTTAATTAATCTGAGAGAG 1572
 1665 TACCTTCAGCTGAGACTTGAACATGAGCCTCGGACAGAGACTTCAAGAAACCGCGGGTGA 1724
 1573 TACCTTCAGCTGAGACTTGAACATGAGCCTCGGACAGAGACTTCAAGAAACCGCGGGTGA 1724
 1725 TTTTGG 1730
 1633 GTGTGG 1638

RESULT 10
 ADR19681
 ID ADR19681 standard; DNA; 1857 BP.
 XX
 AC ADR19681;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human drug metabolizing enzyme (DME)-2 gene sequence.
 XX
 KW drug metabolizing enzyme; DME; cytochrome; immunosuppressive;
 KW anti-inflammatory; endocrine; ophthalmological; gastrointestinal;
 KW hepatocellular; cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; endocrine disorder; eye disorder;
 KW gastrointestinal disorder; liver disorder; metabolic disorder; gene; de;
 human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1.1857
 FT /*tag= a
 FT /product= "Human drug metabolizing enzyme (DME) 2"
 XX
 PN WO200226988-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US030662.
 XX
 PR 29-SEP-2000; 2000US-0236947P.
 PR 06-OCT-2000; 2000US-0238864P.
 PR 20-OCT-2000; 2000US-0242323P.
 PR 09-NOV-2000; 2000US-0247581P.
 PR 16-NOV-2000; 2000US-0249519P.
 PR 22-NOV-2000; 2000US-0252834P.
 PR 30-NOV-2000; 2000US-0250567P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Azimzal Y, Baughin MR, Borowsky ML, Ding L, Dugan BM,
 PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJ, Ison CH, Khan PA;
 PI Lal P, Lee BA, Lu DM, Nguyen DB, Arvizu C, Policky JJ, Rankumner J;
 PI Ring HZ, Sanjamala MS, Tang YF, Tribouley CM, Narinder WK;
 PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 XX
 DR WPI: 2002-362498/39.
 DR P-PSDB; ADR19681.
 XX
 PT Human drug metabolizing enzymes, useful in the diagnosis and treatment of
 PT disorders associated with aberrant (DME) activity, e.g., cancer and
 PT autoimmune disorders.
 XX
 PS Claim 12; SEQ ID NO 20; 142bp; English.
 XX
 CC This invention relates to novel drug metabolizing enzymes (DME) and the
 CC nucleotide sequences which encode them. The invention may be useful for
 CC the development of compounds with a cytochrome, immunosuppressive,
 CC anti-inflammatory, endocrine, ophthalmological, gastrointestinal or
 CC hepatocellular activity acting as an agonist or antagonist of drug
 CC metabolizing enzyme activity. The invention may be used in the diagnosis
 CC and treatment of disorders associated with decreased or increased
 CC expression or activity of drug metabolizing enzymes. Such disorders
 CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
 CC endocrine, eye, gastrointestinal (including liver disorders) and
 CC metabolic disorders. The present sequence is that of a gene which encodes
 CC a human drug metabolizing enzyme (DME) of the invention. Note: This
 CC sequence did not form part of the printed specification but was obtained
 CC in electronic format from EPO.
 XX
 SQ Sequence 1857 BP; 428 A; 487 C; 496 G; 446 T; 0 U; 0 Other;

Query Match

64.2%; Score 1386.2; DB 7; Length 1857;

Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

```
OY 176 GGGAAAACTGGGCGCTTGTGCTGAAGGGCCACAGAGAAACACAGCTGGAGATGATCA 215
Db 176 GGGAAAACTGGGCGCTTGTGCTGAAGGGCCACAGAGAAACACAGCTGGAGATGATCA 215
OY 264 GGTCCAGAGAGGGCTTGTGTGTAAGGGCCACAGAGAAACACAGCTGGAGATGATCA 323
Db 264 GGTCCAGAGAGGGCTTGTGTGTAAGGGCCACAGAGAAACACAGCTGGAGATGATCA 323
OY 236 GGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTTCCGAGTCCC 295
Db 236 GGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTTCCGAGTCCC 295
OY 324 GGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTTCCGAGTCCC 383
Db 324 GGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTTCCGAGTCCC 383
OY 296 CTTTGTGCTCCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTG 355
Db 296 CTTTGTGCTCCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTG 355
OY 384 CTTTGTGCTCCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTG 443
Db 384 CTTTGTGCTCCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTG 443
OY 356 GGAATACCTGGGAGAAAGCCCTCTCACTTATTTGTGCTCCAGAACTCAAGTGGCT 415
Db 356 GGAATACCTGGGAGAAAGCCCTCTCACTTATTTGTGCTCCAGAACTCAAGTGGCT 415
OY 444 GGAATACCTGGGAGAAAGCCCTCTCACTTATTTGTGCTCCAGAACTCAAGTGGCT 503
Db 444 GGAATACCTGGGAGAAAGCCCTCTCACTTATTTGTGCTCCAGAACTCAAGTGGCT 503
OY 416 GCTCTTAGATCAACATGCTCAAGGTGATACCCGAAATTCGAGTGCAGAGACTG 475
Db 416 GCTCTTAGATCAACATGCTCAAGGTGATACCCGAAATTCGAGTGCAGAGACTG 475
OY 504 GCTCTTAGATCAACATGCTCAAGGTGATACCCGAAATTCGAGTGCAGAGACTG 563
Db 504 GCTCTTAGATCAACATGCTCAAGGTGATACCCGAAATTCGAGTGCAGAGACTG 563
OY 476 CCTTAACCTGAACATGCTATGCGCTGCGCCACGCGGATACAGGCTCCAGCTCTT 535
Db 476 CCTTAACCTGAACATGCTATGCGCTGCGCCACGCGGATACAGGCTCCAGCTCTT 535
OY 564 CCTTAACCTGAACATGCTATGCGCTGCGCCACGCGGATACAGGCTCCAGCTCTT 623
Db 564 CCTTAACCTGAACATGCTATGCGCTGCGCCACGCGGATACAGGCTCCAGCTCTT 623
OY 536 GGTGTGTGTTCCAGAGAGTGTGCTTCAAGACTGAGCTCACTTTGATGTGGTCCG 595
Db 536 GGTGTGTGTTCCAGAGAGTGTGCTTCAAGACTGAGCTCACTTTGATGTGGTCCG 595
OY 624 GGTGTGTGTTCCAGAGAGTGTGCTTCAAGACTGAGCTCACTTTGATGTGGTCCG 683
Db 624 GGTGTGTGTTCCAGAGAGTGTGCTTCAAGACTGAGCTCACTTTGATGTGGTCCG 683
OY 596 CCGGTGCTGCTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
Db 596 CCGGTGCTGCTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
OY 684 CCGGTGCTGCTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
Db 684 CCGGTGCTGCTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
OY 656 TTTCTTACACATGAGGATCAGCATCTCGGGAACTGGGCTTCAAGACCAAGTGGC 715
Db 656 TTTCTTACACATGAGGATCAGCATCTCGGGAACTGGGCTTCAAGACCAAGTGGC 715
OY 744 TTTCTTACACATGAGGATCAGCATCTCGGGAACTGGGCTTCAAGACCAAGTGGC 803
Db 744 TTTCTTACACATGAGGATCAGCATCTCGGGAACTGGGCTTCAAGACCAAGTGGC 803
OY 716 TGTCTGTCTGTGGGTCAAGAAACATCGAGTCTTCCGTGGGAGCCCAAGCTGTGAC 775
Db 716 TGTCTGTCTGTGGGTCAAGAAACATCGAGTCTTCCGTGGGAGCCCAAGCTGTGAC 775
OY 804 TGTCTGTCTGTGGGTCAAGAAACATCGAGTCTTCCGTGGGAGCCCAAGCTGTGAC 863
Db 804 TGTCTGTCTGTGGGTCAAGAAACATCGAGTCTTCCGTGGGAGCCCAAGCTGTGAC 863
OY 776 CATCTTGGCGAGTCCGGGGAGCCATAGTGTTCATAGTCTTATCTGTCTCCATGCG 835
Db 776 CATCTTGGCGAGTCCGGGGAGCCATAGTGTTCATAGTCTTATCTGTCTCCATGCG 835
OY 864 CATCTTGGCGAGTCCGGGGAGCCATAGTGTTCATAGTCTTATCTGTCTCCATGCG 923
Db 864 CATCTTGGCGAGTCCGGGGAGCCATAGTGTTCATAGTCTTATCTGTCTCCATGCG 923
OY 836 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCTTACCTGGA 895
Db 836 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCTTACCTGGA 895
OY 924 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCTTACCTGGA 983
Db 924 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCTTACCTGGA 983
OY 896 GGCCCATGATATGAGAAAGTGAAGACCTGCGAGTGTGTCATTTCTGTGTAAACA 955
Db 896 GGCCCATGATATGAGAAAGTGAAGACCTGCGAGTGTGTCATTTCTGTGTAAACA 955
OY 984 GGCCCATGATATGAGAAAGTGAAGACCTGCGAGTGTGTCATTTCTGTGTAAACA 1043
Db 984 GGCCCATGATATGAGAAAGTGAAGACCTGCGAGTGTGTCATTTCTGTGTAAACA 1043
OY 956 TGTGTGAGCTGTGAGGCGCTGTGAGTGTGCTGAGAGAAACCTTCCAGAGAGTGGT 1015
Db 956 TGTGTGAGCTGTGAGGCGCTGTGAGTGTGCTGAGAGAAACCTTCCAGAGAGTGGT 1015
OY 1044 TGTGTGAGCTGTGAGGCGCTGTGAGTGTGCTGAGAGAAACCTTCCAGAGAGTGGT 1103
Db 1044 TGTGTGAGCTGTGAGGCGCTGTGAGTGTGCTGAGAGAAACCTTCCAGAGAGTGGT 1103
OY 1016 GACCCCTCAGCAGAAACAAAGCTTTCATCGAGTGTGATGCTTCTTCTTAA 1075
Db 1016 GACCCCTCAGCAGAAACAAAGCTTTCATCGAGTGTGATGCTTCTTCTTAA 1075
OY 1104 GACCCCTCAGCAGAAACAAAGCTTTCATCGAGTGTGATGCTTCTTCTTAA 1163
Db 1104 GACCCCTCAGCAGAAACAAAGCTTTCATCGAGTGTGATGCTTCTTCTTAA 1163
OY 1076 TGAAGCTCTAGATCTATGCTCTCAGAAAGCATTTAAGCAATTCCTTCATCTGAGT 1135
Db 1076 TGAAGCTCTAGATCTATGCTCTCAGAAAGCATTTAAGCAATTCCTTCATCTGAGT 1135
OY 1164 TGAAGCTCTAGATCTATGCTCTCAGAAAGCATTTAAGCAATTCCTTCATCTGAGT 1223
Db 1164 TGAAGCTCTAGATCTATGCTCTCAGAAAGCATTTAAGCAATTCCTTCATCTGAGT 1223
OY 1136 CAATTAACAAGTGTGCTTCTGTGCTTGAAGAGGCTCCGAGATCTCAAGTGG 1195
Db 1136 CAATTAACAAGTGTGCTTCTGTGCTTGAAGAGGCTCCGAGATCTCAAGTGG 1195
OY 1224 CAATTAACAAGTGTGCTTCTGTGCTTGAAGAGGCTCCGAGATCTCAAGTGG 1255
Db 1224 CAATTAACAAGTGTGCTTCTGTGCTTGAAGAGGCTCCGAGATCTCAAGTGG 1255
OY 1196 CTCACAAGATCCCTTGCCTCATCTGATACAAAACATCTGACACATCCGCGCTCAGTA 1255
Db 1196 CTCACAAGATCCCTTGCCTCATCTGATACAAAACATCTGACACATCCGCGCTCAGTA 1255
```

```
Db 1256 -----TGCATCCCGCTCAGTA 1274
OY 1256 TTTGCACTTGTGGCTAATGAATACCTTCATGACAGACCTCCCTGACGTGAATCCGAG 1315
Db 1275 TTTGCACTTGTGGCTAATGAATACCTTCATGACAGACCTCCCTGACGTGAATCCGAG 1334
OY 1316 CAGTCTTCTGAGACTTGTGAGAGTGTGCTTGTGTGCTGCTGCACTGATCAAGCTCG 1375
Db 1335 CAGTCTTCTGAGACTTGTGAGAGTGTGCTTGTGTGCTGCTGCACTGATCAAGCTCG 1394
OY 1376 ATATCAACAAGATGCTGTGACACTGTCTACTTATGAGATTTCGACACCGGCTCAATG 1435
Db 1395 ATATCAACAAGATGCTGTGACACTGTCTACTTATGAGATTTCGACACCGGCTCAATG 1454
OY 1436 CTTTGAAGACAGAACCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1495
Db 1455 CTTTGAAGACAGAACCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1514
OY 1496 GTTCGGTGTGCTTCTCTGAAGGGGACATTTGATGTTCAAGAGACCAAGAGAGAG 1555
Db 1515 GTTCGGTGTGCTTCTCTGAAGGGGACATTTGATGTTCAAGAGACCAAGAGAGAG 1574
OY 1556 GAAATTACTGAGCCGGAAGATGATGAATATCTGGGCTACCTTTGCTGAAACCGGGAATCC 1615
Db 1575 GAAATTACTGAGCCGGAAGATGATGAATATCTGGGCTACCTTTGCTGAAACCGGGAATCC 1634
OY 1616 TAATGGAAAGACCTGTCTGTGGCAAGCTTATATCTGACTGACAGTACCTCAGCT 1675
Db 1635 TAATGGAAAGACCTGTCTGTGGCAAGCTTATATCTGACTGACAGTACCTCAGCT 1694
OY 1676 GGACTTGAACATGAGCGCTCGACAGAGACTCAAGAAACCGCGGTGATTTTGG 1730
Db 1695 GGACTTGAACATGAGCGCTCGACAGAGACTCAAGAAACCGCGGTGATGTGG 1749

RESULT 11
ADCS5523
ID ADCS5523 standard; cDNA; 1244 BP.
XX
AC ADCS5523;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human carboxylatase 24.64 encoding sequence.
XX
KW human carboxylatase-24.64; primary hypertension; digestive ulcer;
KW nephrotic; bronchial asthma; gene; ss.
XX
OS Homo sapiens.
XX
FT Key Location/Qualifiers
FT CDS 236..910
FT /*tag= a
XX
PN CN1382799-A.
XX
PD 04-DEC-2002.
XX
PF 26-APR-2001; 2001CN-00112736.
XX
PR 26-APR-2001; 2001CN-00112736.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2003-269506/27.
XX
DR P-PDB; ADCS5524.
XX
PT Polypeptide-human carboxylatase-24.64 and polynucleotide for coding it.
XX
PS Claim 6; SEQ ID NO 1; 31bp; Chinese.
XX
```


CC The present invention relates to a polypeptide-human carboxylase-24.64,
CC the polynucleotide for coding it, the process for preparing the
CC polypeptide by DNA recombination, the application of the polypeptide in
CC treating diseases such as primary hypertension, digestive ulcer,
CC nephrotic, bronchial asthma, tremor, etc, the antagonist of the polypeptide
CC and its medical action, and the application of the polynucleotide are
CC new. The present sequence represents human carboxylase 24.64 encoding
CC sequence.
XX
SQ Sequence 1244 BP, 301 A, 314 C, 271 G, 358 T, 0 U, 0 Other;

Query Match 56.3%; Score 1215; DB 10; Length 1244;
Best Local Similarity 99.6%; Pred. No. 1.2e-300;
Matches 1218; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 932 GGTTCACATTTCTGTGTGTAACAATGCTGACACTGAGGCGCTGTGAGTGGCTGAG 991
DB 1 GGTTCACATTTCTGTGTGTAACAATGCTGACACTGAGGCGCTGTGAGTGGCTGAG 60
QY 992 GACAAACCTCCAGAGAGCTGACCTCAGCCAGAGAAACAAGCTTTCAGTCTGAGT 1051
DB 61 GACAAACCTCCAGAGAGCTGACCTCAGCCAGAGAAACAAGCTTTCAGTCTGAGT 120
QY 1052 GGTTCAGTGTCTTTCTTTCTTCTTAATGAGCTTGAATCTATGTCTCAGAAACATTTAA 1111
DB 121 GGTTCAGTGTCTTTCTTTCTTCTTAATGAGCTTGAATCTATGTCTCAGAAACATTTAA 180
QY 1112 AGCAATTCCTTCATCATGAGAGTCAATACACAGAGTGTGGCTTCTGCTGCTTATGAA 1171
DB 181 AGCAATTCCTTCATCATGAGAGTCAATACACAGAGTGTGGCTTCTGCTGCTTATGAA 240
QY 1172 GAGAGCTCTGAGATCTCAGTGGCTCCAAACAAGTCCCTCCTCATCTGATACAAA 1231
DB 241 GAGAGCTCTGAGATCTCAGTGGCTCCAAACAAGTCCCTCCTCATCTGATACAAA 300
QY 1232 CATCTTCGACATCCCGCTCAGTATTTGCACTTTGGCTAATGAATCTTCATGACAA 1291
DB 301 CATCTTCGACATCCCGCTCAGTATTTGCACTTTGGCTAATGAATCTTCATGACAA 360
QY 1292 GCACTCCCTGACATGAAATCCAGAGAGTCTTCTGAGCTTTGGAGATGTTCTTTGT 1351
DB 361 GCACTCCCTGACATGAAATCCAGAGAGTCTTCTGAGCTTTGGAGATGTTCTTTGT 420
QY 1352 GGTCCCTGACATGATCAAGTCTGATATCAACAAGATGCTGTGACCTGTCTACTTCA 1411
DB 421 GGTCCCTGACATGATCAAGTCTGATATCAACAAGATGCTGTGACCTGTCTACTTCA 480
QY 1412 TGAGTTTGGGACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTTGTCAAGCCGA 1471
DB 481 TGAGTTTGGGACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTTGTCAAGCCGA 540
QY 1472 CCACGCTGATGAAGTCCGCTTGTGTTCCGATGCTTCTGAAAGGGGACATTTGTTAT 1531
DB 541 CCACGCTGATGAAGTCCGCTTGTGTTCCGATGCTTCTGAAAGGGGACATTTGTTAT 600
QY 1532 GTTCAAGAGACCGAGAGAGAGAAATTAATGACCGGAAAGATGATGAATTAATCTGAGC 1591
DB 601 GTTCAAGAGACCGAGAGAGAGAAATTAATGACCGGAAAGATGATGAATTAATCTGAGC 660
QY 1592 TACCTTTGCTCGAACCGGGAATCTTAATGAGAACGACCTGTCTGTGGCCAGCTTAATA 1651
DB 661 TACCTTTGCTCGAACCGGGAATCTTAATGAGAACGACCTGTCTGTGGCCAGCTTAATA 720
QY 1652 TCTGACTGACAGTACTCTCAGCTGAACTTGAACATGAGCCTCGAGACAGACTCAAAAG 1711
DB 721 TCTGACTGACAGTACTCTCAGCTGAACTTGAACATGAGCCTCGAGACAGACTCAAAAG 780
QY 1712 ACCGGGGGAGATTTTGGACACAGACATCCCTGATCTGTGTCTGCTCCGACATGCT 1771
DB 781 ACCGGGGGAGATTTTGGACACAGACATCCCTGATCTGTGTCTGCTCCGACATGCT 840
QY 1772 CCACAGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCCAGCCTTTCTTTCTTTT 1831

DB 841 CCACAGTCTCTTTCTTCTTCTTAATCTTCTCTCTCTCTCCAGCCTTCTTTCTTTG 900
QY 1832 TGTCTCTTGAAGAGTATCTTCTGTGATTTGTTTCCCTCTCTCCATTAATTTCTC 1891
DB 901 TGTCTCTTGAAGAGTATCTTCTGTGATTTGTTTCCCTCTCTCCATTAATTTCTC 960
QY 1892 CCGCAATCAATTAATCTTCTTCTGAGCTCAGCTCTTCTATGAGGATCTTGCACAAACA 1951
DB 961 CCGCAATCAATTAATCTTCTTCTGAGCTCAGCTCTTCTATGAGGATCTTGCACAAACA 1020
QY 1952 GGTGCTTCTGATATTTTATGACTTGAAGATGATCTTACAGAAATTTCTTTCAACAT 2011
DB 1021 GGTGCTTCTGATATTTTATGACTTGAAGATGATCTTACAGAAATTTCTTTCAACAT 1080
QY 2012 CAAAAGTGCATTTTGTCTGAGAGCAACAAGATTTCTTCAATTAATTTGAGAGGGC 2071
DB 1081 CAAAAGTGCATTTTGTCTGAGAGCAACAAGATTTCTTCAATTAATTTGAGAGGGC 1140
QY 2072 TGGCTTATTAATGTTCTATTAATGATGTTTGTAACTCATATGAATAATCAGAAATGA 2131
DB 1141 TGGCTTATTAATGTTCTATTAATGATGTTTGTAACTCATATGAATAATCAGAAATGA 1200
QY 2132 AATAGAAAAAATTTTAAAAA 2154
DB 1201 AATAGAAAAAATTTTAAAAA 1223

RESULT 12
ADFS0146
ID ADFS0146 standard; cDNA; 2145 BP.
XX
AC ADFS0146;
XX
DT 12-FEB-2004 (first entry)
XX
DE Cat cauxin encoding cDNA SEQ ID NO:3.
XX
KW cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
XX
OS Felis catus.
XX
PN JP2003250575-A.
XX
PD 09-SEP-2003.
XX
PE 04-MAR-2002; 2002JP-00057908.
XX
PR 04-MAR-2002; 2002JP-00057908.
XX
PA (TOHO-) TOHOKU TECHNORCH KK.
XX
DR WPI; 2004-002277/01.
XX
DR P-PSDB; ADFS0147.
XX
PT Novel cauxin protein or its salt, useful as a cat kidney disease marker.
XX
PS the biological activity of urinary (I) or (VI); and (9) a cauxin
XX
PS Claim 6; SEQ ID NO 3; 33pp; Japanese.
XX
CC The present sequence encodes a cat cauxin protein (I) or its salt, which
CC is cat kidney disease marker. Also described: (1) a partial peptide (II)
CC of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
CC comprising (III); (4) a transformed host (V) comprising (III) or (IV);
CC (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
CC couples specifically with (I) or (II); (7) diagnosing cat kidney disease
CC which involves measuring (I) quantitatively, and where reduction of
CC amount of (I) indicates presence of the disease; (8) a cat kidney disease
CC diagnostic agent comprising (I) labelling agent; a reagent which measures
CC the biological activity of urinary (I) or (VI); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces

CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
 CC blood testing.

XX Sequence 2145 BP; 444 A; 581 C; 550 G; 564 T; 0 U; 6 Other;

Query Match 53.6%; Score 1156; DB 12; Length 2145;
 Best Local Similarity 78.2%; Pred. No. 2.1e-285;
 Matches 1466; Conservative 0; Mismatches 386; Indels 22; Gaps 6;

QY 186 GGGCTTCTGCTGAAGGGCCACAGAGAAACCAAGCTGGATGATTCAGGGCAAGCA 245
 DB 238 GGGCCAGCTGTGATGACACAGAGAGACACAGGAGGATGAGTCCGGGGAGAGCAA 297
 QY 246 GTCACTGTCTGGGAAGCCCTGTGCTGTAAGCTGTCTCTGGAGATCCCTTTGCTGCT 305
 DB 298 ACCACTGTACTGGAGACACCGCTGTGTAACATTTCTCGGAGATCCCTTATGCTGCA 357
 QY 306 CCCCCGCTGGGATCCCTGCGATTTACGAACCCGACAGCTGCATCCGCTGGGATTAATTG 365
 DB 358 CCTCTCTAGGGCCCTCGGATTTAAGCAACAAAGCTGCTGCTGCCCCGGGAATGACTTC 417
 QY 366 CGAGAGCCACCTCTCACTCTAATTTGTGCTTCAGAACTCAGAGTGGCTCTTAGAT 425
 DB 418 CGAAATGCCACATCCCTAACCTTAATATGCTTCCAGAGACTTGAAGTGGCTGCTCTTAT 477
 QY 426 CAACACATGCTCAAGGTCATTACCCGAATTCGAGTGTCAAGAACTGCGCTCTACCTG 485
 DB 478 CAACAGTCTCAAAATGTGCTTAACCCCAATTTGGAAGCTTCGAAACATGCTGTACCTT 537
 QY 486 AACATCTATGCTGCTGCTCCAGCCGATACAGGCTCCAGCTCCGCTCTTGTGTGTTC 545
 DB 538 AACATCTATGCTGCTGCTCCAGCCGATACAGGCTCCAGCTCCGCTCTTGTGTGTTC 597
 QY 546 CCAGAGAGTGTCTTCAAGCTGCTGCTGCTCATTTTGAATGGTTCGCTGCTGCTGCTG 605
 DB 598 CCGGGGGGCTCTTCAAGATGGGCTTCAAGCTTCTCTTGTGATGGGCTCCGCTGCTGCTG 657
 QY 606 TATGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 665
 DB 658 TACGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
 QY 666 ACATGAGATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
 DB 718 ACAGGGGATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
 QY 726 TGGGCTCAGAGAAATCGAGTTCTTGGTGGGAGACCCAGCTCTGTGACATCTTTGGC 785
 DB 778 TGGGCTCAGAGAAATCGAGTTCTTGGTGGGAGACCCAGCTCTGTGACATCTTTGGC 837
 QY 786 GAGTCCGGGGAGCCATAGTGTCTTCTAGTCTTATCTGCTCCCAATGGCCAAAGGCTTA 845
 DB 838 GAGTCCGGGGAGCCATAGTGTCTTCTAGTCTTATCTGCTCCCAATGGCCAAAGGCTTA 897
 QY 846 TTCCACAAAGCCATCATGAGAGAGTGGGAGTGGCCATCATCTTACCTGAG--GCCAT 902
 DB 898 TTCCACAAAGCCATCATGAGAGAGTGGGAGTGGCCATCTGCTTACTGATGAGACCCCT 957
 QY 903 GATTATGAGAGAGTGAAGCTGCTGAGTGGTGGCAATTTCTGTGTGAACAATGCTGCA 962
 DB 958 GGTGATGAGAGAGAGAGATTTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
 QY 963 GACTCTGAGGCGCTGCTGAGTGGCTGAGAGAAACCTCCAAAGAGTGTGAGACCTTC 1022
 DB 1018 GACTCTGAGGCGCTGCTGAGTGGCTGAGAGAAACCTCCAAAGAGTGTGAGACATTC 1077
 QY 1023 AGCCGAAAAACAAGTCTTCACTCGAGTGTGATGTGCTTCTTCTTAATGAGCCT 1082
 DB 1078 AGCAAGAAATCAAGTCTTCACTCGAGTGTGATGTGCTTCTTCTTAATGAGCCT 1137
 QY 1083 CTAGATCTAATGTCTCAGAAAGATTTAAAGCAATTCCTTCAATCATCGAGTCAATTAAC 1142
 DB 1138 GTAGGCTTAATGATCAAAAGCAATTTAATGATTCCTTCTAATCATCGAGTCAATTAAC 1197

QY 1143 CACGAGTGTGCTTCTCTGCTGCTTAAGAGAGGCTCTGAGATCTCTAGTGGCTCAAC 1202
 DB 1198 CACGAGTGTGCTTCTCTGCTGCTTAAGAGAGGCTCTGAGATCTCTAGTGGCTCAAC 1254
 QY 1203 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
 DB 1255 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
 QY 1263 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
 DB 1315 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
 QY 1323 CTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
 DB 1375 CTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434
 QY 1383 AGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1442
 DB 1435 AGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1494
 QY 1443 GACACGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1502
 DB 1495 GACACGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
 QY 1503 GGTGCTTCTCTGAGAGGAGGACATTTGTTATGTTTGAAGAGCCACGAGAGAGAGTA 1562
 DB 1555 GGTGCTTCTCTGAGAGGAGGACATTTGTTATGTTTGAAGAGCCACGAGAGAGAGTA 1614
 QY 1563 CTGAGCGGAGAGATGATGAATTAATCTGAGTACCTTTGCTGAAACCGGAAATCTTA 1622
 DB 1615 CTGAGCGAGAGAGATGATGAATTAATCTGAGTACCTTTGCTGAAACCGGAAATCTTA 1674
 QY 1623 AACAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1682
 DB 1675 GAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1734
 QY 1683 AACATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741
 DB 1735 AAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1794
 QY 1742 -CCCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 DB 1795 GTCCCTGATTAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1854
 QY 1801 TCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
 DB 1855 TTCTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1914
 QY 1861 TTTGGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
 DB 1915 TTGGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1962
 QY 1921 GCTGCTTCTATGAGGATCTTGAACAAACAGCTGCTTGTGCTGATATTTATGACTTA 1980
 DB 1963 CTGTGTTAGTTAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2022
 QY 1981 GGAATGATCTTACAG--AATCTTTTCAACATCAAAAAGTCAATTTGTCTTGAAGGC 2038
 DB 2023 AGGGGATTTTGAAGGGAATTTCTTTCAACCAAAACATCTACCGGCTTTGGAAGGC 2082
 QY 2039 AACAGATTTCTTC 2052
 DB 2083 TACCGATTTCTTC 2096

RESULT 13
 ADF50144
 ID ADF50144 strand: cdna; 1629 BP.
 XX ADF50144;
 AC ADF50144;
 AC ADF50144;
 AC ADF50144;
 DT 12-FEB-2004 (first entry)
 XX

QY 1683 AACATGAGCTCGAGACAGACTCAAGAACCAGGGTGGATTTTGGACAGACCATT 1742
DB 1561 AACTGACGCTGGAGACAGAACTGAAGAGCAAGAGGTGAGTTTGGATGATACATT 1620
QY 1743 CCCC 1746
DB 1621 GTCC 1624
RESULT 14
AB086171
ID AB086171 standard; DNA; 1071 BP.
XX
XX AB086171;
AC 10-SEP-2002 (first entry)
DT
XX
XX Novel human gene. SEQ ID 42.
DE
XX
XX Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antitumor; antitumor; antitumor;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperlipidemia; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo-transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200250105-A1.
PN
XX
XX 27-JUN-2002.
PD
XX 17-DEC-2001; 2001WO-US049232.
PF
XX 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-028135P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAXO) GLAXO GROUP LTD.
XX
XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX MPI: 2002-508784/54.
DR P-PSDB; ABP61006.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX Claim 2(a); Page 252; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,

CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antitumor, antitumor, antitumor, antitumor, antitumor,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records AB086130-AB086184 represent novel human cDNA's
CC of the invention
XX
XX
SO Sequence 1071 BP; 242 A; 286 C; 254 G; 289 T; 0 U; 0 Other;
Query Match 47.3%; Score 1021.4; DB 6; Length 1071;
Best Local Similarity 99.9%; Pred. No. 4,8e-251;
Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 819 ATACGTCTCCCATGAGCCAAAGCTTATTCGCAAGGCATGAGAGTGGGGGCGC 878
DB 49 ATACGTCTCCCATGAGCCAAAGCTTATTCGCAAGGCATGAGAGTGGGGGCGC 108
QY 879 ATACGTCTCCCATGAGCCCAATGATTTAGAAAGTGAAGACCTGAGGTGTTGCA 938
DB 109 ATACGTCTCCCATGAGCCCAATGATTTAGAAAGTGAAGACCTGAGGTGTTGCA 168
QY 939 CATTTCTGTGTACATGCTGACCTGAGGCCCTGCTGAGGTGTTGCA 998
DB 169 CATTTCTGTGTACATGCTGACCTGAGGCCCTGCTGAGGTGTTGCA 228
QY 999 CCTCCAGAGAGCTGCTGAGCCCTGAGCCGCAAGAAACATCTTCACTGAGGTGTTGAT 1058
DB 229 CCTCCAGAGAGCTGCTGAGCCCTGAGCCGCAAGAAACATCTTCACTGAGGTGTTGAT 288
QY 1059 GGTGCTTTCTTCTCTATGAGCCTGATCTATTTGTCTCAGAAAGCATTTAAAGCAAT 1118
DB 289 GGTGCTTTCTTCTCTATGAGCCTGATCTATTTGTCTCAGAAAGCATTTAAAGCAAT 348
QY 1119 CCTTCATCATGAGGTCAATTAACCAAGGTGCTTCTGCTGCTATGAGAGGCT 1178
DB 349 CCTTCATCATGAGGTCAATTAACCAAGGTGCTTCTGCTGCTATGAGAGGCT 408
QY 1179 CCTGAGATCTCAAGTGGCTCCCAAGATCCCTTGCCTTCATCTGATTAACAAACATCCCTG 1238
DB 409 CCTGAGATCTCAAGTGGCTCCCAAGATCCCTTGCCTTCATCTGATTAACAAACATCCCTG 468
QY 1239 CACATCCGCTCAGATTTGACCTTGTGCTAATGAATACCTTCATGACAGCACTCC 1298
DB 469 CACATCCGCTCAGATTTGACCTTGTGCTAATGAATACCTTCATGACAGCACTCC 528
QY 1299 CTGACTGAATTCGAGACAGTCTTCTGACCTTGTGAGATGTCTTGTGTGCTCC 1358
DB 529 CTGACTGAATTCGAGACAGTCTTCTGACCTTGTGAGATGTCTTGTGTGCTCC 588
QY 1359 GCATGATACAGCTCGATATCAAGAGATGTGTGAGCACTGTCTACTTCTATGAGTT 1418
DB 589 GCATGATACAGCTCGATATCAAGAGATGTGTGAGCACTGTCTACTTCTATGAGTT 648
QY 1419 CGGACCGGCTCAGTGTCTTGAAGACAGAGCCGCTTTGTCAAAAGCCAGCAGCT 1478
DB 649 CGGACCGGCTCAGTGTCTTGAAGACAGAGCCGCTTTGTCAAAAGCCAGCAGCT 708
QY 1479 GATGAAGTCGGTTGT 1538
DB 709 GATGAAGTCGGTTGT 768

Db	707	CCATTATTGCACTGTGGCTGTGAATACCTTCATGCAAGCACTCCCTGACTGAAA	766
Qy	1309	TCCGAGACAGCTTTCTGAGCTTGTGAGATGTTCTTTGTGTCCTGCACTGATCA	1368
Db	767	TCCGAGACAGCTTTCTGAGCTTGTGAGATGTTCTTTGTGTCCTGCACTGATCA	826
Qy	1369	CAGCTGATATTCACAGAGATG	1389
Db	827	CAGCTCAGTATCATAGAGTG	847

Search completed: June 14, 2005, 18:10:04
 Job time : 1175.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 7160.54 Seconds

(without alignments)
11471.585 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 2158

Sequence: 1 ccacgcgcgcgaacacag.....aaaaaaaaaaaaaaaaaaaaa 2158

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

BST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_est8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	22.2	742	4	BI822069 603039938
2	418.6	19.4	1680	4	AY414461 Mus muscu
3	414.6	19.2	576	4	BI824830 60303738
4	406.8	18.9	435	4	BM422399 HAS05_70
5	406.6	18.8	1680	4	AY414459 Homo sapi
6	397.8	18.4	2661	3	AK033563 Mus muscu
7	396.2	18.4	2687	3	AK077248 Mus muscu
8	387.6	18.0	4927	3	AK040349 Mus muscu
9	383.6	17.8	1958	3	AK078953 Mus muscu
10	355.2	16.5	1944	3	AK078879 Mus muscu
11	345.8	16.0	2888	3	HSMB03013
12	344.2	15.9	3909	3	HSMB06270
13	338	15.7	1918	7	CF111083
14	333.8	15.5	989	3	AK007235
15	331.4	15.4	2038	3	BC019926 Mus muscu
16	298.4	13.8	1826	3	CR618303 full-len
17	293.6	13.6	1859	3	CR641368 Tetradon
18	292	13.5	1817	3	CR634885 Tetradon
19	289	13.4	1510	3	AK050444 Mus muscu
20	286	13.3	2356	3	AK037191 Mus muscu
21	285	13.2	491	5	BX280446 BX280446
22	284.2	13.2	1849	5	CR647187 Tetradon
23	284	13.2	539	5	BUS85267
24	284	13.2	593	1	AI808985

25	282.2	13.1	805	5	BP442124	BP442124
26	282.2	13.1	808	5	BP446194	BP446194
27	282.2	13.1	840	5	BP445396	BP445396
28	282.2	13.1	841	5	BP447764	BP447764
29	282.2	13.1	860	5	BP441209	BP441209
30	282	13.1	904	5	BQ900930	BQ900930
31	281.8	13.1	798	5	BP444116	BP444116
32	281.8	13.1	808	5	BP443157	BP443157
33	281.8	13.1	877	5	BP441564	BP441564
34	281.8	13.1	877	5	BP442855	BP442855
35	281.6	13.0	1823	3	CR640788	CR640788
36	280.8	13.0	1853	3	CR723405	CR723405
37	280.2	13.0	840	5	BP442111	BP442111
38	279.8	13.0	783	5	BP445505	BP445505
39	279.6	13.0	792	5	BP443197	BP443197
40	276	12.8	645	5	BUS85270	BUS85270
41	275.4	12.8	905	7	CO583389	CO583389
42	274	12.7	1857	3	CR637401	CR637401
43	273.6	12.7	1117	7	CN061747	CN061747
44	272.4	12.6	763	7	CR765061	CR765061
45	272	12.6	770	5	BP444251	BP444251

ALIGNMENTS

RESULT 1
BI822069
LOCUS 603039938F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181056 5', mRNA sequence.
DEFINITION BI822069
ACCESSION BI822069.1 GI:15933619
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Place: LLM11451 row: k column: 09
High quality sequence stop: 740.
Location/Qualifiers
1..742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5181056"
/lab_host="DH10B"
/clone_lib="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NOT; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 22.2%; Score 479; DB 4; Length 742;
Best Local Similarity 100.0%; Pred. No. 9.1e-119;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 GGGCTTTGCTGAAAGGCGCACAGAGAACACAGGCTGGATGATTCAGGCGAACAA 245
DB 103 GGGCTTTGCTGAAAGGCGCACAGAGAACACAGGCTGGATGATTCAGGCGAACAA 162
QY 246 GTCACTGTGCTGGAAAGGCGCGCTGTGTGAAGCTGTTCTCGGAGTCCCTTTGCTGCT 305
DB 163 GTCACTGTGCTGGAAAGGCGCGCTGTGTGAAGCTGTTCTCGGAGTCCCTTTGCTGCT 222
QY 306 CCCCCTGGGATCCCTGCGATTACGAACCCGAGCTGATGCCCTGGAAATCTTG 355
DB 223 CCCCCTGGGATCCCTGCGATTACGAACCCGAGCTGATGCCCTGGAAATCTTG 282
QY 366 CGAAGACCGACCTCTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTCTTAGAT 425
DB 283 CGAAGACCGACCTCTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTCTTAGAT 342
QY 426 CAACACATGCTCAAGGTCATTAACCGAAATTCGAGATGTCAGAACTGCTTACCTG 485
DB 343 CAACACATGCTCAAGGTCATTAACCGAAATTCGAGATGTCAGAACTGCTTACCTG 402
QY 486 AACATCATGCGCTGCTCCAGCGCGATACAGGCTCCAGCTCCCTTTGTGTGTTTC 545
DB 403 AACATCATGCGCTGCTCCAGCGCGATACAGGCTCCAGCTCCCTTTGTGTGTTTC 462
QY 546 CCAAGAGTGTCTTCAAGACTGCTCAGCTCCATCTTGTAGAGGTCCGCTGCTGCTC 605
DB 463 CCAAGAGTGTCTTCAAGACTGCTCAGCTCCATCTTGTAGAGGTCCGCTGCTGCTC 522
QY 606 TATGAGAGCTGTGTTGTGTGCTGCTCAGTACCGGCTAGAAATTTGTTTCTTCA 664
DB 523 TATGAGAGCTGTGTTGTGTGCTGCTCAGTACCGGCTAGAAATTTGTTTCTTCA 581

RESULT 2
AY414461 1680 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus CES2 gene, VIRTUAL TRANSCRIPT, partial sequence.
DEFINITION AY414461
ACCESSION AY414461 GI:3970423
VERSION GSS.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1680)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL PUBMED 14671302
REFERENCE 2 (bases 1 to 1680)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
SOURCE 1..1680
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

gene <1..>1680
/gene="CES2"
/locus_tag="HCM5231"

Query Match 19.4%; Score 418.6; DB 9; Length 1680;
Best Local Similarity 56.9%; Pred. No. 3.5e-102;
Matches 885; Conservative 0; Mismatches 634; Indels 36; Gaps 5;

ORIGIN

QY 210 AGGACACACAGCTGGATGATTCAGGCGAACAGTCACTGTCTGGAGGACCTGTG 269
DB 100 AGGACACACACAGCTGGATGATTCAGGCGAACAGTCACTGTCTGGAGGACCTGTG 159
QY 270 CCGTGAACGTGTTCTCTGAGATCCCTTTGCTGCTCCCGCGGATCCCTGCAATT 329
DB 160 GGTGTCCACTTTCTCTGGGATCCCTTTGCTGAGCTCTGTAGAACACTGCGCTTT 219
QY 330 AGGAACCCGACGCTGCATCCCTGGGATTACTTGGAGAGCCACCTCTAACCCTAAT 389
DB 220 GCACCCCTGAGCCACTGAACCATGGGGGTGTGTGAGAGTGGACATCATCCGGCC 279
QY 390 TTGTGCTCCAG---AACTGAGTGTGCTGCTTATGATCAACATGCTCAAGCTGAT 446
DB 280 ATGTGTGAGAGCATCATGCAATATATGAGAGCTTTTAAGCTTTGAAGCTGAC 339
QY 447 TACCGAAATTCGAGTGTCAAGACTGCTCTAACCCTGAACATTAAGCCCTGAC 506
DB 340 TTGCTCTCCCTCTAATGTCTGAGACTGCTCTAATCTTAACATTAATGACCAATCAT 399
QY 507 GCCGATACAGGCTCAAGCTCCCGCTTGTGTGTGTGTTCCAGAGTGTCTTCAAGCT 566
DB 400 GCCGATACAGGCTCAAGCTCCCGCTTGTGTGTGTGTTCCAGAGTGTCTTCAAGCT 459
QY 567 GGTCAAGCTCACTTTGATGAGTCCGCTGCTGCTCTATGAGAGCGTGTGTTG 626
DB 460 GGTCAAGCTCACTTTGATGAGTCCGCTGCTGCTCTATGAGAGCGTGTGTTG 519
QY 627 GTGTCACATACCGGCTAGGAATTTGTTTCTTCAACATGAGATGAGATGCTCCG 686
DB 520 ACTATCCAGTACCGGCTAGGAATTTGTTTCTTCAACATGAGATGAGATGCTCCG 579
QY 687 GGGACGTGGGCTTCAAGAGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
DB 580 GGGACGTGGGCTTCAAGAGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639
QY 747 TTTTGTGTTGGGAGCCCACTGTGACCATTTTGGGAGTCCGCGAGCCATAAGT 806
DB 640 CACTTTGTGTTGGGAGCCCACTGTGACCATTTTGGGAGTCCGCGAGCCATAAGT 699
QY 807 GTTCTAGTCTTAATCTGCTCCCAATGGCCAAAGCTTATTCACAAAGCCATCTGAG 866
DB 700 GTTCTAGTCTTAATCTGCTCCCAATGGCCAAAGCTTATTCACAAAGCCATCTGAG 759
QY 867 AGTGGGTGGGATCATCTTACCTGAGAGGCCATGATTAAGAGAGAGTGAAGACTG 926
DB 760 AGTGGGTGGGATCATCTTACCTGAGAGGCCATGATTAAGAGAGAGTGAAGACTG 813
QY 927 CAGGTGTGACATTTTCTGTGTAACAATGCTCAGACTCTGAGGCCCTGCTGAGTGC 986
DB 814 ACTTACGTGTGACATTTTCTGTGTAACAATGCTCAGACTCTGAGGCCCTGCTGAGTGC 873
QY 987 CTGAGACAAACCTCCAAAGAGCTGTGACCTGAGCCGAGAAACAAAGTCTTTCACT 1046
DB 874 CTGAGACAAAGAGTGAAGAGAGATTTGTGCTTATTAAGAGATTTCCAGATATCCCT 933
QY 1047 CGAGTGTGATGCTGCTTTTCTTCTAATAGGCTGTAGATCTTATGTCTCAGAAAGCA 1106
DB 934 GGTGTGTGATGAGTGAAGTCTTACCAAGATCTCTCAGAGGCTGTGAGCTGTGCTGAT 993
QY 1107 TTAAAGCAATTCCTCATCATGAGTCAATTAACAGAGTGTGCTTCTGCTGCT 1166
DB 994 TTCACTGTGCTTCCAGATCATTTGTTTCAACATGATGATGATGATGCTTCC 1053


```

Qy 1167 -----ATGAGAGAGGCTCTGAGATCTCAGTGGCTCCAGAAAGTCCCTTGGC 1214
Db 1054 AAAGCATGAGCATATGATGCTCCAGAAAGGAGTGAACAGCAATGGTGCAGCTATC 1113
Qy 1215 CTCCTCTGATACAAACATCTCGACATCCCGCTCAGTATTTGACCTTGGCTAAT 1274
Db 1114 CTGCAGAGAGGCAACCGATGATGCTGCTCTCTGATGATGACCTGTCTAATGGA 1173
Qy 1275 GAATACCTTCATGACAGACCTCCCTGACTGAAATCCGAGACGCTCTGAGCTTGGCTT 1334
Db 1174 GAGTACATGGGGGACCTGAGAGATGCCAGACCTCCCAATACAGTTTCAAGAGATGATG 1233
Qy 1335 GAGATGATGTTCTTTGATGCTCCCTGCACTGATCAGAGCTGATCAGAGAGATCTGGT 1394
Db 1234 GAGATTTTCAATGTTTGTATGCTTCCATCTCCAGATGACATTTTCAGAG---TTCCAT 1290
Qy 1395 GCACTGTCTACTTCTATGATGTTTGGGACCGGCTCAGTGTCTTGAAGACAGAGCCG 1454
Db 1291 GCGCTGTCTACTTCTATGATGTTTCAACATCAATCAGCTTCTCAAGATGTCAAGCC 1350
Qy 1455 GCTTTTTCAGAAAGCGACAGCTGATGAAAGTCCGCTTTGTTGGTGGTCCCTTCTG 1514
Db 1351 CCCACGTAAGCGCTGACCATGAGATGAGTTCCTTTGTTTGGTCTTCTTCTG 1410
Qy 1515 AAGGGGACATTTGATGTTGAGAGAGCAGAGAGAGAGATTACTGAGCCGAG 1574
Db 1411 GGCATGAATCTTGACTT-----CACTGAGAGAGAGAGAGAGCTGTGAGAGAG 1458
Qy 1575 ATGATGAATACTGGGCTACCTTTGCTGAAACCGGAAATCTTAATGGAACGACTGTCT 1634
Db 1459 ATGATGAATGACTGGGCAACTTTGCAAGACATGAGAAACCCCAAGTGAAGGCTTACCC 1518
Qy 1635 CTGAGGCAAGCTTAATCTGACAGAGTACCTCCAGCTGGAATTTGAACATGAGCCTC 1694
Db 1519 TACTGAGCTTGATGAGACATGATGAGATGCTGACAGCTGGAATTCAGACTTCTGTG 1578
Qy 1695 GAGACAGAGACTCAAGAACCGGAGTGAATTTTGAACAGACCATCCCTGTA 1749
Db 1579 GGTGAGGCTTGAAGGCGAGAGGCTGCACTTCTGACCAAGACTTGTGCCAG 1633

```

```

RESULT 3
LOCUS B1824830 576 bp mRNA linear EST 04-OCT-2001
DEFINITION 60303373871 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174937 5',
mRNA sequence.
ACCESSION B1824830
VERSION B1824830.1 GI:15936380
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 576)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: L14M11435 row: 1 column: 10
High quality sequence start: 6
High quality sequence stop: 565.
Location/Qualifiers

```

```

FEATURES
SOURCE
1..576
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:5174937"
/lab_host="DH10B"
/clone_id="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC library."

```

```

Query Match 19.2%; Score 414.6; DB 4; Length 576;
Best Local Similarity 96.4%; Pred. No. 3e-101;
Matches 456; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

```

```

Qy 186 GGGCTTGTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 245
Db 104 GGGCTTGTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 163
Qy 246 GTCACTGTGCTGGAAGGCTGTGCTGTGAACGTGTTCTCGAGTCCCTT-TGCTGC 304
Db 164 GTCACTGTGCTGGAAGGCTGTGCTGTGAACGTGTTCTCGAGTCCCTTATGCTGC 223
Qy 305 TCCCCCGTGGGATCCCTGCGATTTACGAACCCGACGCTGATCGCCCTGGATTACTT 364
Db 224 TCCCCCGTGGGATCCCTGCGATTTACGAACCCGACGCTGATCGCCCTGGGATTACTT 283
Qy 365 GCGAAGAACCCACCTTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTTGA 424
Db 284 GCGAAGAACCCACCTTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTTGA 343
Qy 425 TCAACACATGCTCAAGGTGATTAACCCGAATTTGGG-AGTGTCAAGAACCTGCTTACC 483
Db 344 TCAACACATGCTCAAGGTGATTAACCCGAATTTGGGAGTGTCAAGAACCTGCTTACC 403
Qy 484 TGAACATGATGCGCTGCTGCCACAGCCGATACAGGCTCCAGCTCCCTTGTGTGTGT 543
Db 404 TGAACATGATGCGCTGCTGCCACAGCCGATACAGGCTCCAGCTCCCTTGTGTGTGT 463
Qy 544 TCCACAGAGTGGCTTCAAGACTGCTGCTGATCCATCTTTGATGAGTGGTCCGCTGGCTG 603
Db 464 TCCACAGAGTGGCTTCAAGACTGCTGATCCATCTTTGATGAGTGGTCCGCTGGCTG 522
Qy 604 CCTATGAGACAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
Db 523 CTATGAGACAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575

```

```

RESULT 4
LOCUS BM422399 435 bp mRNA linear EST 01-FEB-2003
DEFINITION HAS05_70 Human testis cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION BM422399
VERSION BM422399.1 GI:28191734
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 435)
Liang, G., Miao, S.-Y., Zhang, X.-D., Wang, L.-J. and Wang, L.-F.
The differentially expressed gene in spermatogenesis
Unpublished (2002)
COMMENT
Contact: Lin-Fang Wang
National Laboratory of Medical Molecular Biology
Institute of Basic Medical Sciences, Chinese Academy of Medical
Sciences & Peking Union Medical College; Chinese National Human

```

Genome Center, Beijing
#5, Dong Dan 3 Tiao, Beijing 100005, P.R.China; #3-707 North
Yongchang Road BDA, Beijing 100176, P.R.China
Tel: 8610-65396418
Fax: 8610-65240529
Email: wangli@cdm.lmicams.ac.cn.

FEATURES
Source location/Qualifiers

1..435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/clone_lib="Human testis cDNA library"
/note="The cells of Primary Spermatoocyte and Round
Spermatid were isolated from Human testis. Total RNA was
extracted by Trizol. Using PCR-Select cDNA Subtraction
Kit, the library was constructed"

ORIGIN

Query Match 18.9%; Score 406.8; DB 4; Length 435;
Best Local Similarity 99.5%; Pred. No. 3.7e-99;
Matches 408; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1666 ACCTCCAGCTGGAAGTGAACATGAGCTCGAGACAGAGACTCAAGAAACCGGGTGGATT 1725
DB 1 ACCTCCAGCTGGAAGTGAACATGAGCTCGAGACAGAGACTCAAGAAACCGGGTGGATT 60
QY 1726 TTGGACACGACACATCCCCCTGATCTGTCACCTCCGACATGCTCCAGTCTCTTT 1785
DB 61 TTGGACACGACACATCCCCCTGATCTGTCACCTCCGACATGCTCCAGTCTCTTT 120
QY 1786 CTTCCTTAACTTCTCTCTCTCTCTCCACGCTTCTTTCTTTTGTGTGCTCTTGAAG 1845
DB 121 CTTCCTTAACTTCTCTCTCTCTCTCCACGCTTCTTTCTTTTGTGTGCTCTTGAAG 180
QY 1846 TTATCTTCTGATTTTGTGTTTCCCTTCTCTCCCATTAATTTCTCCGCAATCTTAGC 1905
DB 181 TTATCTTCTGATTTTGTGTTTCCCTTCTCTCCCATTAATTTCTCCGCAATCTTAGC 240
QY 1906 TTCTTCTGAGCTCAGCTCTTTCTATGGGAGATCCCTGCAAAAGAGCTTTCGCTGA 1965
DB 241 TTCTTCTGAGCTCAGCTCTTTCTATGGGAGATCCCTGCAAAAGAGCTTTCGCTGA 300
QY 1966 TATTTTATGAGACTTGAATGATCTTACAGAAATCTTTTCAACATCAAAAAGTCAATT 2025
DB 301 TATTTTATGAGACTTGAATGATCTTACAGAAATCTTTTCAACATCAAAAAGTCAATT 360
QY 2026 TGTCTTGAAGGACACAGAAATTTCTTCAATTAATTTGGAAGAGGCTGGC 2075
DB 361 TGTCTTGAAGGACACAGAAATTTCTTCAATTAATTTGGAAGAGGCTGGC 410

RESULT 5
AY14459 1680 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens CES2 gene, VIRIDAL TRANSCRIPT, partial sequence,
DEFINITION Genomic survey sequence.
ACCESSION AY14459
VERSION AY14459.1 GI:39770421
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1680)
AUTHORS Clark,A.G., Glanowski,S., Nelson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1680)
AUTHORS Clark,A.G., Glanowski,S., Nelson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission
COMMENT Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source location/Qualifiers
1..1680
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/gene="CES2"
/locus_tag="HCM5231"

ORIGIN

Query Match 18.8%; Score 406.6; DB 9; Length 1680;
Best Local Similarity 56.5%; Pred. No. 6.5e-99;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;
QY 211 GGAACACGAGCTGGAGATGATTCAGGGCAAGCACTGTCGGGAAGCGCTGTGC 270
DB 101 GGACACACACACAGGGGCAAGTGTGGAGTCTTGTTCATATGAAAGGCGCAATGCC 160
QY 271 CTGTGAACGTTCTCTCGAGTCCCTTGTCTCTCTCCCGCTGGATCTTCGATTTA 330
DB 161 GGGTCAACCTTCTCGGAAATTCATTGGCAAGCACCTTAGGTCCGCTCGCATTTG 220
QY 331 CGAACCCGACCTGTGATGCGCTGGGATTAATTTGCAAGACACCTCTACCTTAAT 390
DB 221 CACCCCTGAGCCCGCTGAATCTTGAATGTGTATGAGATGAAACCACTTCGGCCA 280
QY 391 TGTGCTCCGAACTCAGATGAGTCTCTTGA--TCACACATGCTCAAGTGCATT 447
DB 281 TGTGCTCCGAACTCAGATGAGTCTCTTGA--TCACACATGCTCAAGTGCATT 340
QY 448 ACCGAAATTTGGAGTGTGAGAGATGCTCTACCTGAACATTAATGCGCTGCCACG 507
DB 341 TCCCTTCGACATCAATGTCTGAGATGCTCTTACCTCAGCATTCACACGCGGCCATA 400
QY 508 CCGATACAGGCTCCAAAGTCTCCCGCTTGTGTGTTCCAGAGAGTCTTCAAGATG 567
DB 401 GCCATGAAGGCTCTTAACCTGCGGTGATGTGTGATCAAGGTGTGCGCTGTTTGG 460
QY 568 GCTCAGCTCCATCTTTATATGAGTCCGCTGAGTCTGCTATGAGACGTGCTGTGG 627
DB 461 GATAGCTTCTTGTATATGATGTTCATGCTGCTGCTTGAAGAACCTGTGTGTGCA 520
QY 628 TGTCCAGTACCGGCTAGAAATTTGTTTCTTACACATGAGGATACAGATCTCCG 687
DB 521 TCATCAAGTACCGGCTGAGTGTCTGAGCTTCTTACAGACTGAGACAGACCAACCG 580
QY 688 GGAACCTGGGCTTCAAGAACCAAGTGTCTGTCTGTGGTCCAGAAACATTCAGT 747
DB 581 GCAACCTGGGCTTCAAGAACCAAGTGTCTGTCTGTGGTCCAGAAACATTCAGT 640
QY 748 TCTTGGTGGGACCCCAAGCTCTGTGACCATCTTTGGGAGTCCGGGGGCCATAAGT 807
DB 641 ACTTTGAGGGAACCTTGAACCTGTACCATTTTGGCGAGTCTGCGGGGCGACGAGT 700
QY 808 TTTCTAGCTTAATCTGTCTTCCATGCGCAAGGCTTAATTCACAAAGCATATGAGA 867
DB 701 TGTCTTGGCTTGTGTGCTCCCAATATCCCAAGGACTTTCACAGGAGCATATGAGA 760
QY 868 GTGGGAGTGCATATCCCTTACTGAGGCGCATATTAATGAAAGTGAAGACCTGC 927
DB 761 GTGGGAGTGCATATCCCTTACTGAGGCGCATATTAATGAAAGTGAAGACCTGC 814

QY 928 AGGTGTTGCACTTTCTGTGTGTAACAATGCGTCAGACTGTGAGGCCGCTGAGGTGCC 987
DB 815 CGGTGTGTGCAACTCTGTCTGTGCTGTGACCAAGTTGACTGTGAGGCCCTGTGTGGCTGCC 874
QY 988 TGAGGACAAAACCTCTCAAGAGAGCTGTGACCTTCAGCCAGAAAACAAAGTCTTTCCTC 1047
DB 875 TGCGGGGCAAGATGTAAGAGAGATTTCTGCAATTAAGAGCTTTCAGATGATCCCG 934
QY 1048 GAGTGTGTGATGTGTCTTCTTCTTAATGAGCTGTGATGTATTTGTCTCAGAAAGCAT 1107
DB 935 GAGTGTGTGATGTGTGTCTTCTTCTGCGCAGAGCACTCCAGAGAGTGTGTGCTGTGCCACT 994
QY 1108 TTAAGCAAAATCTCTCAATCAATCGAGTCAATACAGAGTGTGTCTCTGCTGCTCA 1167
DB 995 TTCAAGCTGTCTCTGCAATTTGTGTGTCAACAATGAAATTCGGCTGTGCTATCCCA 1054
QY 1168 TGA-----AGGAGGCTCTGAGATCTCTGAGTCTGCAAGTCCCTTGTGCC 1215
DB 1055 AGGTGATGAGATCTATGATACCAAGAAAGAAATGAGACAGAGGCTCCAGAGTGTCTC 1114
QY 1216 TCCATCTGATACAAACAATCTGTGCAATCCGCTCAGTATTTGCACTTGTGTGCTAATG 1275
DB 1115 TGCAAAAATGTAAAGCTGTGATGTGTGCTCTCAATTTGTGTGACCTGTGTGAGGAGG 1174
QY 1276 AATACCTTCATGACAGAGCACTCCCTGAGTGAATCCGAGACAGTCTTGTGAGCTTGTG 1335
DB 1175 AGTACATTTGGGACATGAGGAGATCCCAAGACCTTCAGAGCGAGTTTCAAGAGATGATG 1234
QY 1336 GAGATGTGTCTTGTGTGCTCTGCACTGATCAAGCTGATGATCAAGAGATGTGTG 1395
DB 1235 CGGACTCATGTGTGTGTGATCTCTGCACTCAAGTACATTTTCAAGT---TTCGCGG 1291
QY 1396 CACCTGTCTACTTCTATGAGTGTGCGCACCGGCTCAGTGTCTTGAAGACACAGACCGG 1455
DB 1292 CCCTGTGTACTTCTACAGAGTTCAGAGATCAGCCAGCTGAGTCAAGAAATCAGGCGAC 1351
QY 1456 CTTTGTGCAAGCCGACACAGCTGATGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTG 1515
DB 1352 CGACATGAAAGGACACATGATGATGAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTG 1411
QY 1516 AGGGGACATTTGTTTGTTCAGAGAGCCAGAGAGAGAGTACTGAGCCGAGAA 1575
DB 1412 GCACTACTATTAAT-----CAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1459
QY 1576 TGATGAAATCTGAGCTACTTGTCTGCAACCGGAAATCTTAATGGAACGACTGTCTC 1635
DB 1460 TGATGAAATCTGAGGCAATTTGCGAAGAAATGGAAACCCCAATGGCGAGGTCTGCCAC 1519
QY 1636 TGTGTGCGCAGCTTAATATCTGACTGAGAGTACTCTGAGTGTGAAATGAACTGAGCTCG 1695
DB 1520 ACTGTGCGCTGTTCAGACAGAGAGAGAGCAATCTGAGTGAACCTACAGCTGTGCGTGTG 1579
QY 1696 GACAGAGACTCAAGAACCGCGGCTGAGATTTTGGACCGACCATCCGCC 1746
DB 1580 GCCGGGCTCTGAGGCGCCAGAGGCTCCAGTTCTGGAAGAGCGCTGCCCC 1630

RESULT 6
AK033563
LOCUS
DEFINITION
AK033563 2661 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male colon cDNA, Riken full-length inserted
library, clone:9030624L02 product:similar to CARBOXYL-ESTERASE
PRECURSOR (EC 3.1.1.1) (AL-ESTERASE) (B-ESTERASE) (MONOBUYRASE)
(COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
(Mesocricetus auratus), full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK033563.1 GI:26329244
HTC; CAP trapper.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nageoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsumai, T., Teshiro, H., Itoh, A.,
Sunt, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2661)
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirokane, T.,
Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE

JOURNAL
Submitted (16-JUL-2001) Yoshinori Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-research@riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers

FEATURES

Source
1..2661
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:9030624L02"

```

/db xref="taxon:10090"
/clone="9030624L02"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
25..1704
/notes="unnamed protein product; putative
similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)
(ALT-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE
ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
[meocricetus auratus] (SPTR|O35533, evidence: FASTY,
83%ID, 100%length, match=1677)"
/codon_start=1
/protein_id="BAC28361.1"
/db xref="GI:26329245"
/translation="MPYKILGMLNANVACGVLLLVHVGQDSASPIRHTTQVGS
LVHYKDDIADVHTPLGIPFAKPPVPLRFAPPEAPPEPMSGVDRGTSHPMCLQNDMLM
GSEDLKMNLLLPISMEDCLYLNIVYPHAHGSNLPMVWHGGALITGVASMYD
GSMILATEDVVVAIQRGLVGFSTGDHAKGNMGYLDPOVALRMVONITVREGN
PDRVLTIGESAGTSVSHVSPMSQGLFGLIMESGVATLPLISSSEMHRIYAN
LSGCAVNSEETLMCKLKEAEMLAINIKFKIIPGVQSEPIPKRPOELMASKDPH
VPSIIGINDEYGMILPTIMDPAOKIEIRKTLPAVLSTALKMLPEECDDLME
YMDPTDEPILQAPREMKDPMFVLPALDVAFHORSHPVVEFEQHRSPSEKFRP
PYKADHDEITFLVFGTQFQGNITLPTVEEEOQLSRMMKRYANFAHGHNPSEGLPYW
PWNHDEOYLQDIPQSVGRALARRIQFWTKTLPKQIKELKSGRHKEL"
2634..2639
/notes="putative"
2661
/notes="putative"

```

```

polya_signal
polya_site
ORIGIN

```

```

Query Match      18.4%; Score 397.8; DB 3; Length 2661;
Best Local Similarity 56.1%; Pred. No. 1.9e-96;
Matches 872; Conservative 0; Mismatches 647; Indels 36; Gaps 5;

210 AGGAACACCAAGCTGGATGATTCAGGGCAGCAAGTCACTGTCTGGAAAGCCCTGTG 269
Db      124 AGGAACACACACAGCCAGGAGAGAGCAGCTTGTCCAGTAAAGACCTGACATC 183
QY      270 CCTGTACAGTGTCTCTGGAAGTCCCTTTGTGTCTCCCGCTGGAGATCCCTGGCATTT 329
Db      184 GCTGTTCATACCTTCTCGGGAATCCCTTTGCGCAAGCCCTGTGAGACCACTGGGCTTT 243
QY      330 AGCAACCCGACGCTGCATCGCCCTGGAGTAATCTTCCAGAAAGCACTTCTTACCTTAAT 389
Db      244 GACCTCTCGAGAGCCCTGAGCCATGAGTGGGGTGAAGATGGGATCTCAATCAAAAC 303
QY      390 TTGTGCTTCAGAACTCAGAGTGTGCTCTTA--GATCAACACATCTCAAGGTGAT 446
Db      304 ATGTGCTTACAAATGACAAATTGTATGGGTTCAAGAGATCTGMAAGATATTAATGATC 363
QY      447 TAACCGGAATTCGAGGTCTGAGAAAGTCTCTTACCTTAACATCTATCGCCCTGCCAC 506
Db      364 CTGCTCCCATCTTCTATGTCTGAGAGACTGCTGTATCTCAACATCTATGTATCCAGGCCAT 423
QY      507 GCCGATACAGGCTCCAGCTCCCGCTTGTGTGTGTGTTCCAGAGAGGTGCTTCAAGACT 566
Db      424 GCCCATGAGGGTTCTAATCTTACCTGTATGATGATGATTCATGATGATGATGATGATGAT 483
QY      567 GGCCTACGCTCTTCTTGTATGAGTCCGCTGCTGCTCTATGAGAGAGTGTGTTGTG 626
Db      484 GGCATGCTTCTCATGTATGATGATTCATGCTGCAACAATGAGATGTGTGTGTGTG 543
QY      627 GTCTGTCAAGTACCGGCTAGGAATTTGTGTTCTTTCACATGAGGATTCAGATGCTCCG 686
Db      544 GCTATTCAGATACCGCTCGGGTGTCTTGGAAATCTTTCAGACTGGAAGCCAGCATGCGAAA 603
QY      687 GGAACCTGGGCTTCAAGAACCAAGTGTCTGTCTGTCTGGGTCCAGAAAGCATTCGAG 746
Db      604 GGGAACTGGGGATACCTGGAACCAAGTGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 663
QY      747 TTCTTGGTGGGAGCCCAAGCTCTGTGACATCTTTGGCGAGTCCGCGGAGGACATTAAGT 806

```

```

Db      664 CACTTGGAGGCAACCTTCAACGATGACATTTTGGAGTCCGCGTGGCAAAAGT 723
QY      807 GTTTCATGCTTATACCTGTCTCCATGCGCAAAAGCTTATTCACAAAGCATCATGAG 866
Db      724 GTGTCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 783
QY      867 AGTGGGTGGCCATCATCCCTTACCTGAGAGGCCATGATTTAGAGAGTGAAGACCTG 926
Db      784 AGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 837
QY      927 CAGGTGTGTCACATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 986
Db      838 AGAATAGTGGCCAAATCTATCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
QY      987 CTGAGAGCAAAACCTCCAGAGAGCTGTGACCTCAGCCAGAAACAAAGTCTTTCAT 1046
Db      898 CTAAAGAGCAAAATGAAGAGAGATCTGTCTTAAATAGGCTTCAAAATCATCTCT 957
QY      1047 CGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1106
Db      958 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
QY      1107 TTTAAAGCAATCTCTTCAATCATGAGTCAATTAACCAAGTGTGTGTGTGTGTGTGTGTGT 1166
Db      1018 TTTCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1077
QY      1167 ATGAAGAGGCTCTCT-----GAGATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
Db      1078 ACATCATGAGCCAGCTCAGAAATATGAGAAATTAACAGAAAGCCCTGCAAGCTGT 1137
QY      1215 CTTCATGTATTAACAAACATCTGTGACATCTCCGCTCAATTTTGTGACCTGTGTGTGT 1274
Db      1138 CTAAAGAGCACACCTTAATAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1197
QY      1275 GAATACTTCCATGACAGACAGCTCCCTGACATGAATCCGAGACAGTCTTGTGACTTGTCT 1334
Db      1198 GAGTACATGGGGGACACTGAGAGCCAGAGACTGTGCAAGACACAGTTCAAGAGATGAAG 1257
QY      1335 GAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1394
Db      1258 GGGAGCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1314
QY      1395 GCACGTGTCTATCTTATGATGTTTGGGACCGGCTCAGTGTCTTGAAGACAGAGCCG 1454
Db      1315 GCTTCTGTCTTATCTTATGATGTTTCAACATGACACCACTTCTTCAAGAGATTTGAGGCCA 1374
QY      1455 GCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGT 1514
Db      1375 CCTATGTGAAGCTGACATGATGTATGAATTTTCTTGTCTTGTGTGTGTGTGTGTGTGT 1434
QY      1515 AAGGGGACATTTGTATGTGTGAAGAGCCACGAGAGAGAAATTATCTGAGCCGGAAG 1574
Db      1435 AACATTAACCTTCTTA-----CACTAGAGAGAGAGCAACTTAAGCAAGAG 1482
QY      1575 ATGATGAATACTGGGCTTACCTTGTGTGAAACGGGAATCTTAATGGAAAGACCTGTCT 1634
Db      1483 ATGATGAAGTACCGGCTCACTTGTGACAGGAGGAAACCCCAACATGTAGGGGTCTAACCC 1542
QY      1635 CTGTGGCCAGCTTATATCTGATGAGAGTACCTTCAAGCTGTGACATGAGGCTTCTGTG 1694
Db      1543 TACTGGCTGTGATGACATGATGACAGATGACATGACATGACATGACATGACATGACATG 1602
QY      1695 GGAACAGACTCAAGAAACCGCGGGTGAATTTTGTGACAGACACATCTCCCTCTGA 1749
Db      1603 GGTGAGGCTTGAAGGCAAGAGCTCAAGTCTGAGCAAGACTGTGCCCAGA 1657

```

```

RESULT 7
AK077248
LOCUS
DEFINITION
AK077248 2687 bp mRNA linear HTC 03-APR-2004
Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
RIKEN full-length enriched library, clone:503145B19

```


D	512	GGATAGGCTTCATGTATATGATGATTCATGTGGCAGCACTGAGAGATGTGTGTGTC	571
Q	627	GTCTGCAGTACCGGCTAGAGAAATATTTGGTTCTTCAACAATGGAGATCAGACATGCC	686
D	572	GCATTCACATACCGCTGGGTGTCTGGGATTTCTTACGACTCGAGACACACATGCCAA	6311
Q	687	GGGAACTGGGACCTTCAAGGACCAAGTGTGTGTCTGTCTTGGGTCCAGAAACATTCAG	746
D	632	GGCAATGGGGATACCTGGACCAATGTGTGTCTGTCTGTGGGTCCAGCAAAACATTTGC	6911
Q	747	TTCTTGGTGGGACCCCACTGTGTGACCAATCTTTGGCGAGTCCGGGGAGCCATTAAT	806
D	692	CACTTTGGGAGCAACCTTGACCGAGTCAACAATTTTGGAGGTACAGAGGTGGCAAGT	751
Q	807	GTTCCTAGCTTATATCTGTCTTCCATGGCCAAAGCTTATCCAAAGCCATCATGGAG	866
D	752	GTGTCTTCAATGTGTGTCTCCCATGTGCCAGGAGACTTCTCCATGGTGCATCATGGAG	811
Q	867	AGTGGGAGTGGCCATCATCTTACCTGGAGGCCCATGATTAATGAAAGATGAGACCTG	926
D	812	AGTGAAGTAAAGCTGTGTCTCTGTATCTTATCTCCAGCTCTCTGAGA-----TGTTCAC	865
Q	927	CAGGTGTGTGACATTTCTGTGTGTAACAATGCGTCAACTCTGAGGCCCTGTGAGGTGC	986
D	866	AGATAGTGGCCAACTTATCTGTGTGTGACAGCTGTGAACTCAAGACCCGTATGTCTGC	925
Q	987	CTGAGACAAAACCCCTCCAGAGAGCTGTACCCCTGACGAGAAAACAAGTCTTCACT	1046
D	926	CTAAGAGGCAAGATGAAAGAGAGATCTGGCTTATTAAGGTCTTCAAAATCAATCCT	985
Q	1047	CGATGTGTATAGTGTCTTTCTTCTTAATGAGCCTCTAGATCTATTTGTCTGAGAAAGA	1106
D	986	GGTGTGTGTGATGAGAGTTCACCCAGACATCTCAAGAGCTGATGGCTCTAAGAT	1045
Q	1107	TTTAAAGCAATTCCTTCATCATCTGAGAGTCAATPAACAAGATGTGGCTTCTGTCTCT	1166
D	1046	TTTCAACCTGTCCCAAGCATATTGGATTCACAAATGAAATATGTGTGATTTCTCT	1105
Q	1167	ATGAGGAGGCTCT-----GAGATCTCAGTGTCTCCAAAGTCCCTTCC	1214
D	1106	ACGATCATGAGACCCAGACTCAAAATATAGAGAAATPAACAAGAAACCTGTGACGTCT	1165
Q	1215	CTCCATCTGTATCAAAACAATCCGCAATCCGCGCTCAAGTATTTGCACTTGTGTCTAT	1274
D	1166	CTAAAAAGCAGCCCTTAAAAATGATGTCTCTCTGATGTGTGACCTTCTATGAG	1225
Q	1275	GATACTCTCATGACAAGACATCTCCCTGACTGAAATCGAGACAGTCTTGTGACTTCT	1334
D	1226	GAGTATATGAGGGGACCTGAGAGCCAGAGACTGTGCAAGCAGTTCAAGAGATTAAG	1285
Q	1335	GGAGATGTGTCTTGTGTGTGTCTCCCTGCACTGATCAGAGTCAAGTCAAGAGATGTGT	1394
D	1286	GGGAGCTTATGTGTGTGTATCTCTGTGACTTCAAGTATGACATTTTCA---ACGTTCCAT	1342
Q	1395	GCACTGTCTACTTCTATGAGTTTGGCAGCGGCTCAGTGCTTTGGAAGAACAAGACCG	1454
D	1343	GCTCTGTCTACTTCTATGAGTTCACATCATGACCCAGCTTCTTCAAGATTTCAAGCCA	1402
Q	1455	GCTTTGTCAAAAGCGACACGCGTGAATATCGGCTTGTGTGTGTGTGTGTGTGTGTGT	1514
D	1403	CCCTATGTGAAGGCTGACCATGTGTATTAATTTCTGTCTTTGTGCTACAGTTTCGT	1462
Q	1515	AAAGGGGACATTTGTATGTGAAAGAGCCAGAGAGAGAGAAATTACTGAGCCGAGAG	1574
D	1463	AAACATAAACCTCTTA-----CACTAGGAAAGAGAGCACTTAACAGAAAG	1510
Q	1575	ATGATGAATACTGGGCTACTTGTCTGAAACGGGAACTCTAATGGGAAAGCATGTCT	1634
D	1511	ATATATGAATGTACTGGGCACTTTGCAACGGCAGGAAACCCCAACAGTGTGCTTACCC	1570
Q	1635	CTGTGGCAGCTTAAATCTGACGACAGTAACTCAGCTGAGCTTGAACATAGACCTTC	1694
D	1571	TACTGTGCTGTGATGAGCAATGATGAGACATGTAAGTGTGACATTCAGCTTCTGTG	1630

QY		1695	GACAGACGACTCAAGAACCCTGGGTGATTTTGGACACGACCATCCCTCGA	1749
Dd		1631	GGTCGAGCCTTAAGGCCAGAAAGGCTGCAGTTCGACCAAGACTCGCCACGA	1685
RESULT 8				
LOCUS	AK040349			
DEFINITION	AK040349 Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library; clone:A430088E12 product:similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1) (AL1-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYL-BUTYRASE) [Mesocricetus auratus], full insert sequence.			
ACCESSION	AK040349			
VERSION	AK040349.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Bukacynska, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P., and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Nature Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komno, H., Atiyama, J., Niishi, K., Kitsumura, T., Tashiro, H., Itoh, M., Suni, L., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwaig, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiru, A., and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076861			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
REFERENCE	5			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
JOURNAL	Nature 420, 563-573 (2002)			
REFERENCE	6			
AUTHORS	(bases 1 to 4927)			
TITLE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroyuki, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawaji, J., Kohjima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, N., Nihi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saio, R., Satoh, K., Sekai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takeku-Akahira, S., Takeuchi, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.			
REFERENCE	Direct Submission			

JOURNAL

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/; Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

polyA_site

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.

URL: http://genome.gsc.riken.jp/

URL: http://phantom.gsc.riken.jp/

URL: http://phantom.qualifiers

FEATURES

source

1. 4927

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:A430088E12"

/db_xref="taxon:10090"

/clone="A430088E12"

/cissue_type="chymus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

25. 1711

/note="putative

similar to CARBOXYLESTERASE PRECURSOR (EC 3.1.1.1)

(ALI-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE

ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)

[Mesocricetus auratus] (SPTR103553, evidence: PASTV,

70.7kD, 99.2%length, match=1672)"

4912. 4917

/note="putative"

4927

/note="putative"

ORIGIN

Query Match 18.0%; Score 387.6; DB 3; Length 4927;

Best Local Similarity 56.1%; Pred. No. 1.4e-93;

Matches 890; Conservative 0; Mismatches 659; Indels 37; Gaps 7;

185 TGGGCTCTGCTGGAAGGCGCAGAGAACCAAGGCTGGATGATTCAGGCGAAGCA 244

105 TGACTCATGAGGCGCAGCCCATGAAACACACTTCTGACAGGTCGAGGCAAGTT 164

245 AGTCACTGTGCTGGAAGCCCTGTGCTGTGAACGTGTTCTGAGAGTCCCTTTGGTC 304

165 TGTCACTTGAAGACATCAAGCTGAGGCGCATTAATTCTGGGGATCCCTTTGCCAA 224

305 TCCCCCGCTGGATCCCTGCGATTAGAAC-CCGAGCTTCGATCCCTTGGGATTA 363

225 GCTCTCTGTGAGACCATGCGCTTGTGACCTCTGAGGCGCCCTGAGCATGAGTGTG 284

364 TGGGAGAGCCACTCTTACCTTAATTGTGCTCAGAACTCAGAGTGGCTCTTAG 423

285 TGAAGAGTGGACCTCAGCGCGCCATGTGTAAATAATGATGATGATGAAATTAG 344

424 A---TCACACATGCTCAAGTGTACCCGAAATTCGAGTGTCAAGAGCTGCTCT 480

345 AGGCTGTAAGATTAATAAGATGATCTGCTCTCTCTTCTATGTCTGAGACTGCTCT 404

481 ACTTGAACATCTTATGCGCTGCGCAGCGGATACAGGCTCAAGCTCCCGTCTTGCT 540

405 ATCTCAACATCTTACACACAGCCAGCCAGAGGCTCTTAACCTGCTGATGATGT 464

541 GGTTCACGAGAGTCTTCAAGACTGCTCAGCTTCAGCTCTTATGATGATGATGATG 600

465 GGTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 524

601 CTGCTTATGAGAGCTGTGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

525 CAGCCACTGAGATGT 584

661 TCACCAATGAGATGT 720

585 ACAGCACTGAGATGT 644

721 TGTCTGAGT 780

645 TACACTGAGT 704

781 TTGGCGAGT 840

705 TTGGCGAGT 764

841 GCTTATTCACAAAGCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

765 GACTCTTCATGAGCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 824

901 ATGATTTATGAG 960

825 ACTCCCTGAGATGA-----TCTTCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 878

961 CAGACTGTGAG 1020

879 CAACTCAGAAAGCCCTGT 938

1021 TCAGCCAGAAACAAAGT 1080

939 TTAGCAAGGCTTCAAGT 998

1081 CTCTAGATCTATGT 1140

999 CCAAGAGT 1058

1141 ACCAGAGT 1188

1059 ATGATGAGATGAGT 1118

1189 TCAGT 1248

1119 TAAATGAGAAATGATGT 1178

1249 CTCAATTTGTGACCTGT 1308

1179 CTGATGT 1238

1309 TCGAGACAGT 1368

1239 TCAATGACATGT 1298

1369 CAGCTGT 1428

1299 TTAGCAGT 1355

1429 CTGAGT 1488

1356 CAGT 1415

1489 GCTTGT 1548

1416 ACTTATCTT---TGGATCTCTTCTGTGGGCTCAAAATTATGCTT-----CACTG 1463

1549 AGGAGAGAGATGT 1608

1464 CCGAGAGAGAGT 1523

1609 GGAATCTTAAAG 1668

1524 GGAACCCCAAGT 1583

1669 TCCAGCTGT 1728

1584 TGCAGCTGT 1643

QY 1729 GGACGACGACATCCCGTCGATCCG 1754
|||||
Db 1644 GGACCAAGACTCTTCCCGCAAGATG 1669

RESULT 9
AK078953
LOCUS 1958 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130231C15 product:similar to LIVER CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus], full insert sequence.
AK078953
VERSION AK078953.1 GI:26347654
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasahagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861

REFERENCE
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12051100

REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, A., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of

COMMENT
Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>.
URL: <http://fantom.gsc.riken.jp/>.
Location/Qualifiers
1. 1958
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM.DB:9130231C15"
/db_xref="taxon:10090"
/clone="9130231C15"
/sex="male"
/tissue_type="cecum"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
66. 1742
/note="unnamed protein product; putative similar to LIVER CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus] (SWISSPROT|064419, evidence: FASTY, 78.4%id, 100%length, match=1674)"
/codon_start=1
/protein_id="BAC37476.1"
/db_xref="GI:26347655"
/translation="MPLRLPGMLCVVACGLLLIQLHYHGDSAPIRNTHRGVRS FVNVQDKGVAFGLIPIPAKPPVGLRPARPEDEPMSGYRDTGSOAPMLQDIPYN LBEAKMNLIPISMBEDCLYNTTPHQAQESNLVYMWIHGGILVGSASANDV SKLAATEEIVIVAIQVRLGIVGFPSTGOHARGNGVITDOVALKRWKNTAVFGNR DRYITGVSAIGGATVSHTLSFMSKGLFHGALMOSGVALLDLSIDSEVYKTVANL SCGEANDSALHCHLRPAKSKOSTLAINOVFNKIPAVVDGEFLPKIPDILTSMFVAVL PRTIGVNTDECGVPMFMGDLHIIPKNTRETLPAVLKNTAARMKLPESGHLVVEY WEDTDPEPTLQVPEFEMLDPEFVLPALQVAFQPSQAPVAFYEFQHLSPFKVPS HYKADHGDAVFVGSYLDNMNLDTEEEELKRMKMYMANPARNGNSEGLPSNP VLDHBOYLQDLPQPAVGRALKARLQFWTTLTPQKIQELKSGODKRAEL"
1937. 1942
/note="putative"
1958
/note="putative"
polyA_site
polyA_site
ORIGIN
Query Match 17.8%; Score 383.6; DB 3; Length 1958;
Best Local Similarity 55.1%; Pred. No. 1.3e-92;
Matches 855; Conservative 0; Mismatches 664; Indels 33; Gaps 4;
QY 210 AGGAACACGAGCTGGATGATTCAGGCAAGCAAGTCACTGCTGGAGAGCCCTGTG 269
|||||
Db 165 AGGAACACCAAGGAGGTCAGAGTGAAGAGCGAGCTTGTGCCAGTGAAGACATCAATCT 224
|||||
QY 270 CCGTGAACGATGTCCTGGAGTCCCTTGGCTGCTCCCGCGTGGATCCCTGGCATTT 329
|||||
Db 225 GGTGTCATGCTTCTGGGAATCCCTTTGGCAAGCTCTCTGAGGACTGATCGCGCTTT 284
|||||
QY 330 AGGAACCGGACGCTGATGCGCTGGAGATTAATCTGGAGAAGCCACTCTCACTCAAT 389
|||||
Db 285 GCACCTCTGTGAGCACTGTAGCCATGAGTGGTGTGAGATGGAGCTCAGACCGGCC 344
|||||
QY 390 TTGTGCTCCGAGACTGAGTGGCTGCTTGTAGTCAACAATGCTCAAGTGATAC 449
|||||
Db 345 ATGTGCTGACGCTGATTAATTAATTTAGAGATGCGAAGAAATGATCGATCTG 404
|||||
QY 450 CCGAAATTCGAGATGAGAGAGCTGCTTACCTGAACATTAATGCGCTGCGACGGC 509
|||||
Db 405 CTTCCACTCTATGTCGAGAGCTGCTGTATCTCAACATCTACACCAACCAACCGCC 464
|||||

Qy	510	GATACAGGCTCCAAAGCTCCCGCTCTTGAGTGTGTTCCAGAGAGGTGCTTCAAGACGTGC	563
Db	465	CAGAGGGCTCTAACTCTGCTGTGATGATGATTCACCGGTGGTGACTGTGTGTAGAA	524
Qy	570	TCAGGCTCCATCTTGTATGATGGGTCCCGCTGGGTGCTATAGAGACGTGCTGGTGTGGCT	629
Db	525	TCAGGCTTCATGAATGATGATGTATCTAACTGGCAGCCTAGAGAAATAGTATTTGTGC	584
Qy	630	GTCAGTACCGGATAGAAATATTTGGTTCTTTCACACATGGGATCAGCATGCTCCGGG	689
Db	585	ATCCAGTATGCTGTGGGGTCTTGCTTCTTTCAGACCTGGAGACCAAGCATGCCAGAGGC	644
Qy	690	AACGTGGCTTCAGAGACCAAGTGGCTGCTGTGCTCTGGGTTCAGAAACAATGAGTTC	749
Db	645	AACGTGGGATACCTGGACCAAGTGGCTCTTCACTGGGTGTCAGAAATAATGCTTAC	704
Qy	750	TTCCGTTGGGGACCCGACGCTGTGTACATCTTTGGCGAGTCCGGGGGACATAAGTGT	809
Db	705	TTTGAGGCAACCGTGAACGGGTCACTAATTTGGAGTCTAAGCAGGAGGTACAAGTGTG	764
Qy	810	TCATGCTTATACTGTCTTCCATGCGCCAAAGGCTTATTCACAAAGCCATCATGAGAGT	869
Db	765	TCTTCTCATATTTCTGTCCCAATGTCCAAAGGACTTTCAATGAGGCATTATGAGAGT	824
Qy	870	GGGTGGCCATCATCTCTTAACTGGAGGCCATGATATAGAGAGTGAAGACCTGAG	929
Db	825	GGAATGGCCCTGTGCTGCTGACCTCATCTTGTACACTTCTGAG-----GTGGTCTACAG	878
Qy	930	GTGGTGTGCAATTTCTGTGTAAACAATGGTCAGCTGTGAGGCCCTGTAGAGGCTGTG	989
Db	879	ACAGTGGCAACTTATCTGTGTGTGAACCCAGACTGAAGAGCCCTATATCACTGCTGTG	938
Qy	990	AGGACAAACCTCCAGAGAGCTGTGACCTCAGCCAGAAACAAAGTCTTTCACCTGCA	1049
Db	939	AGAGAAAGATGAACAAGAGATCTGTGCTATTAACAGAGTCTTCAATATATCCCTGCT	998
Qy	1050	GTGGTGTGCTGCTTCTTCTTCTAATAGCCCTAGATCTATGTCTCAGAAAGCATTT	1109
Db	999	GTGGTGTATGGGAGTTTCTTACCCAGCATCTTCAGAGCTGTGACCTTATGATTTT	1058
Qy	1110	AAAGCAATTCCTTCATCATCGGAGTCAATAACACAGTGTGGTTCCTGTGCTATAG	1165
Db	1059	CACCTGTGCCCCAGCATATGTGTGTCAACAAGATGATGTGGTGGAGTCCCAATG	1118
Qy	1170	AAGAGGCTCCCG-----AGATCCGACGTGCTCCAAACAAATCCCTTGCCCTC	1217
Db	1119	TTTCATGGGCTTGTACATATCATTAAGAAATATACAGAGAGCCCTGTCCAGCTGTCTG	1178
Qy	1218	CATCTGATACAAACATCTGTGCACATCCCGCTCAGTATTTTGGACCTTGTGGCTAATGAA	1277
Db	1179	AAGAAACACAGACGACGAGATATGTGTCTCCAGAAATGATGTACCTGTAGTGGAGAG	1238
Qy	1278	TACTTCATGACAGACACTCCCTGTGCTGAATTCGAGACAGTCTTGTGACTTGTGGA	1337
Db	1239	TATATGGGGGACACTGAGAGACCCAGAGACCCTGTCAAGACAGTTCAAGAAATGTCTGGG	1298
Qy	1338	GATGTGTTCTTTGTGTGCTCCCTGTGCACGTATACAGCTGTATATCAAGAGATGCTGTGCA	1397
Db	1299	GACTTCATGTTGTGTATCCCTGTGCT--TCAAGTGGACAAATTTCCAGCGTTCCAGGCT	1355
Qy	1398	CTGTGCTACTTCTATGAGTTTTCGCAACGGGCTCAGTGTCTTTTGAAGACCAAGAGCGGCT	1457
Db	1356	CTGTGCTACTTCTATGAGTTTTCAGATCTGTGTCCAGCTTCAACAGCAATGTACAGGCTCA	1415
Qy	1458	TTTGTCAAGCCGACACGCTGTATGAGTCCGCTTTGTGTGTGGTGTGCTTCTCTGAAG	1517
Db	1416	CATGTGAAGGACAGACCATGTGTATGATGACCTTGTGTGTGGCTCTATATTTGTGGAC	1475
Qy	1518	GGGACATTTGTATGTGGAAGAGCCACGAGAGAGGAAGTTATCTAGCCGGAAGATG	1577
Db	1476	ATGATTTTGGACT-----CACTAGAGAGAGAGGCTGCTAAGAGATGATGTG	1523

QY	1578	ATGAATAACGCGGCTACTCTTGGCTCGAACCGGGAATCTTATGGGAACGACCTGTCTG	1637
Db	1524	ATGAAGTACTGGGCGCAATTTCGCAAGAAATGGGAACCCCAACTATGAGGGCTTAACTCC	1583
QY	1638	TGGCAGAGTTATATCTGACTGAGCAGTACCTCCAGCTGAGCTTGAACATGAGCCTCGA	1697
Db	1584	TGGCTCTGTGTGGACCATGATGACAGTAACTGCAAGCTGGACACCCAGCCTGCTGTGGGT	1643
QY	1698	CAGAGACTCAAGAACCAGCGGGTGGATTTTGTGACCAAGACATATCCCTCTGA	1749
Db	1644	CGAGCCTTGAAAGGCCAAGAGGCTCAAGTTCTGGAACCAAGACTCTGCCCCAGA	1695

LOCUS	AK078879	1944 bp	mRNA	linear	HTC 03-APR-2004
DEFINITION	Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030604P03 product:carboxylesterase 3, full insert sequence.				
ACCESSION	AK078879				
VERSION	AK078879.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
REFERENCE	1				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komio, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P., Komio, H., Akiyama, J., Nishii, K., Kikunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530513				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5				
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.				
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs				
JOURNAL	Nature 420, 563-573 (2002)				
REFERENCE	6 (bases 1 to 1944)				
AUTHORS	Aach, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Hayashizaki, S., Furum, M., Hanagaki, T., Hara, A., Hashizume, W., Kiyukida, S., Kurumoto, N., Hiramoto, K., Hirose, T., Hirozane, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komio, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Osato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,				

TITLE
JOURNAL

Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akhiba, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

COMMENT

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gscl.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES
Source

1. 1944
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:9030604P03"
/db_xref="taxon:10090"
/clone="9030604P03"
/sex="male"
/clone_type="colony"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
64..1761
/note="unlabeled protein product; carboxylesterase 3 (MGI:2148202, GB|NW_053200, evidence: BLASTN, 99%, match=1939)
putative"
/codon_start=1
/protein_id="BAC37439.1"
/db_xref="GI:26347581"
/translation="MGLYPLIMLSIACTAMGYPSPVNVTVKGVLGKYNLEGPQVAVFLGVPAKPLPLSLRFAFPOPAKPSFVKNTSYPMCSGDVAGGVLSLEPTNKENIPLQPSDECLYNIYTPADLTNRSLPVMWHIGGLVVGASTYGLSLSHENVVVTIOYRLGIMGFSTGDSHRISGWGHIIDVAAIRWQDNIANGSGPSVTFHGSAGFSVSVLWSPILAKNLPHRAISSGVSLTAALITTVKPIAGLVATLSCGKTTLSAIVWHICRQTEDELEISLKLNLFLDLGNPKESYPLPTVIGVVLSPKAEI LAKSPSTVPTIVGINKQEPFWIIPITLMGYPLABGKLDQKRNALLMKSYPYTKISEN MIPVAEKYLGSTDLTKKKDLFDLMDVVGVPVSIVSRSHDAGASTYMEFEYR PSFVAMRPKAVIGDHGDEIFSVFSGPLKDGASEETNLISKVMKFMANPARENGNPGGLPHMPEYDOKEGYLKIAGSTQAQRLKQKVSFMAELRAKESAPRSHREVEL"
1922..1927
/note="putative"
1944
/note="putative"

CDS

polyA_signal
polyA_site
ORIGIN
1944
/note="putative"

Query Match 16.5%; Score 355.2; DB 3; Length 1944;
Best Local Similarity 55.1%; Pred. No. 7e-85;

Matches 870; Conservative 0; Mismatches 633; Indels 75; Gaps 6;

QY 212 GAACACCGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGGCC 271
Db 141 GAACACCTGTTAAAGCAAGTCTGGGGAGTAGTCAATTGGAGGATTCACACAGCC 200
QY 272 TGTGAAGCTGTTCTCTGGAGTCCCTTGTGCTGCTCCCGCTGGATCCCTGCGATTAC 331
Db 201 TGTGGCTGTTTCTCTGGAGTCCCTTGTGCGAAGCCCTCTTGGCTCTTGAAGATTGC 260
QY 332 GAACCGGACGCTGCGATCGCTCTGGATTAATTGCGAAGACCACTCTCTAATTT 391
Db 261 TCACACCAACGCTGCAAGCCCTGGAGCTTGTGAAGAAACACACCTCTCAACCGCTAT 320
QY 392 GTGCTCTCAAGCTC-----AAGTGGCTGCTTATGATCAACACATGCTCAAGGTGA 445

Db 321 GTGCTCTCAAGATGCTTTGGAGGGAGAGTGTCTCAAGCTCTTACCAACAGAAAGA 380
QY 446 TTAACCGAAATTCGAGTGTCAAGAACTGCTCTTACCTGAAACATCTATGCGCTGCCA 505
Db 381 GAACATTCCTTACAGATTTTTCAGAAAGCTGCTTACCTGAATATTTACATCTCTGCA 440
QY 506 CCGCGATACAGGCTCCAAAGCTCCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565
Db 441 CTTGACAAAGAACAGAC 500
QY 566 TGGCTCAGCTTCATCTTGTATGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625
Db 501 GGGCGAGACATCACTATGATGAGCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
QY 626 GGTGCTCAGTACCGGCTAAGGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 685
Db 561 GACCTTATGATGAGCTTGGCATCTGGGATTCCTTACGACCTGGGGATGACACAGTGC 620
QY 686 GGGGAACCTGGGCTTCAAGACACAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 745
Db 621 GGGAAACTGGGGTCACTTGGACAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
QY 746 GTTCTGCTGGGAGACCCAGCTGTGTGACCATCTTGTGGAGTCCGCGGAGCCATAAG 805
Db 681 CAATCTTGGGGGCAACCCAGCTGTGTGACCATCTTGTGGAGTCCGCGGAGTTCAG 740
QY 806 TGTCTTATGCTTATATCTGTCTCCATGAGCAAGGCTTATCCAAAGCCATCATGA 865
Db 741 TGTCTGCTTCTGT 800
QY 866 GAGTGGGCTGCTCATCTTCTTACCTGAGGCCATGATTTAGAAAGTGAAGACCT 925
Db 801 GAGTGGTGTCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
QY 926 GCAAGTGTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
Db 861 GATGCTACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
QY 986 CTTGAGACA-----AAACCTCCAAAGAGCTG 1013
Db 921 CCGAAAGACAGAGATGAACCTAGAGACCTCACTAAATTTGAATCTTTTAACTGA 980
QY 1014 CTGACCTCAGCGAAGAAACAAG-----TCTTCACTGAGTGTGTGTGTGTGTGTGTGT 1066
Db 981 CTTACTTGAATTCGAAGAGAGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1040
QY 1067 CTTTCTTAATGAGCTCTAGATCTATGCTCTAGAAACATTTTAAAGCAATTCCTTCAT 1126
Db 1041 TCTGCCAAAGGACACAGAGATCTGTGAGAAAGTTTTCAGACTGTCTCTCTCTCTCTCT 1100
QY 1127 CATCGAGTCAATTAACACAGATGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1177
Db 1101 AGTGGCATCAACAGAGAGTGTGTGCTGATCTTCAACGCTTATGAGCTATTCACCT 1160
QY 1178 TCTTGAGATCTCACTGAGTGTCTCAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
Db 1161 CGCTGAAGCAACCTGAGCAGAAAGACGCAATCTCTCTTGTGGAAGTCTTACCCAC 1220
QY 1235 CCTGACATCTCCGCTCAGTATTTGCACTTGTGCTATGAATTAATCTCATGACAAACA 1294
Db 1221 ACTTAAATCTCTGAGATATGATTTCCAGTGTGTGCTGAGAAAGTATTTAGAGAGACAG 1280
QY 1295 CTCCCTGAGTAATTCGAGACAGCTTCTGAGCTTCTGAGATGTGTGTGTGTGTGTGTGTGT 1354
Db 1281 TGACCTACCAAAAAGAAAGACCTTCTCAGACCTGATGTGTGTGTGTGTGTGTGTGTGTGT 1340
QY 1355 CCTGCACTGATCAACAGCTGATATCAAGAGATGTGTGTGCACTGTCTATCTTATGA 1414
Db 1341 CCATCATGATGT 1400
QY 1415 GTTTCGGACCGGGCTCAAGTCTTGAAGACAGAAAGCGGCTTTTGTCAAAAGCCAGCA 1474

Db 1401 GTTTAGATATCCGCCAGCTTTGTATTCGGCCATGAGACCCAGCAAGTAAATAGAGACCA 1460
Qy 1475 CGCTGATGAAGTCCGCTTTGTGTGGTGGTGGCTCTTCTGAAAGGGGACATTGTTATGTT 1534
Db 1461 TGGTATGAGATCTTCTCACTATTTGATCTCCATTTTAA----- 1502
Qy 1535 CGAAGACCCAGAGAGAGAGAGTAACTGAGCCGAGAGATGAAATACTGGGCTTAC 1594
Db 1503 AGATGGTCTCCAGAAAGAGAGACCACTGACGAAGATGTGATGAATTCCTGGGCCAA 1562
Qy 1555 CTTTCTCGAACCAGGGAATCTTAATGGAACGCTGTCTGTGGCCAGCTTAATATCT 1654
Db 1563 CTTTCTCGAATGGAACCCCAATGTGAGAGGCTGCCCCCACTGGCCGAATATGACCA 1622
Qy 1655 GACTGAGAGTACCTCCAGCTGATGAAACATGAGCCTCGACAGAGACTCAAGAAC 1714
Db 1633 GAGGAGGGGTATCTGAAAGATTGTGCTCACTCAAGGACCCAGAGGCTTAAGAGCAA 1682
Qy 1715 GCGGGTGAATTTTGGAC 1732
Db 1683 AGAAGTGAATTTTGGGC 1700

RESULT 11
HSM803013 2888 bp mRNA linear HTC 22-SEP-2004
LOCUS HSM803013
DEFINITION Homo sapiens mRNA; cDNA DKFZp434N0935 (from clone DKFZp434N0935).
ACCESSION AL713761
VERSION AL713761.1 GI:19584506
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2888)
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weil, B., Amid, C., Oeinger, A., Fobo, G., Han, M. and
Wiemann, S.
COMMENT The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neubiberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp434N0935) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneId=DKFZp434N0935
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/
Location/Qualifiers
1..2888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434N0935"
/db_xref="taxon:9606"
/clone="DKFZp434N0935"
/cfeature_type="CDS"
/clone_id="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="cardoxylesterase 2 isoform 2"
1..2888
/gene="DKFZp434N0935"
954..2777
/gene="DKFZp434N0935"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD28531.1"
/db_xref="GI:19584507"
/db_xref="UniProt/Swiss-Prot:O00748"

/translation="MTAQRSPPTPTFPSPQRTPLNCPVQTPRLGALIHCTWDPG
QPLAQGRVRRQRTETSSPTWRLHRLRLRLAVACGLLLVRCGGQDSASPIRTHT
GOVGSILVHVGAAGVOTFLGIPAKPLPGLRPAKPSGSGVSGVGLDTPMLPCL
ODLTAVERSEPLSOPFMTPSPDSMSRDCILSYLTPHAKHBSNLVPMWIIHGAIVG
MASLYDGSMLALENVVVITITGLVGLFSPSTGKHTKNGVYLDVLAALRYVOON
AHFGNPDRTVITPESAGTSVSLVSPISQGLPHGAIMESGVALPEGLIASADVI
STVANILASAOVQSEALVGLRGSKSEIILAINPFOIIPGVVGLPRPISILA
SADFQVPSIYGVNNNEFMGLPKVMRIYDIOKEMDREASQALQKMLTLPPTFG
DLUREYIGDNDPOTLOAOFQEMMADSMFYI PALOVAHFOCSRAPIYFVRFQPSW
LKNIRPMKADHYVTFEERDLSRQMKVYANFARNPNNEGIPHPPLFQBEQOYL
QNLTPAVGRALKHRLQPKKALPQKIQELFEERHRTTEL"

ORIGIN
Query Match 16.0%; Score 345.8; DB 3; Length 2888;
Best Local Similarity 54.8%; Pred. No. 2.9e-82;
Matches 850; Conservative 0; Mismatches 617; Indels 84; Gaps 5;

Qy 211 GGACACCAAGCTGGAGATGATTCAGGGCAAGCAAGTCACTGTCTGGAGACCTGTGTC 270
Db 1246 GGACACACACACAGGGGAGGAGTGTGGGGAAGTCTTGCATGATGAAGGCGCCAAATGGCG 1305
Qy 271 CTGTGAAGCTGTTCTCGGAGTCCCTTTGCTGCTCCCCGGTGGATCCCTGCATTTA 330
Db 1306 GGGTCCAAACCTTCTGGGAATTCATTGTCAGAGCCACTCTAGAGTCCGCTGCAATTTG 1365
Qy 331 CGAACCCGAGCTGCATCGCCCTGGAGTAACTTGCAGAAACCACTCTCACTCAATT 390
Db 1366 CACCCCTGAGGCCCTCTGAATTTGAGATGTGTGAGGATGGAACCACTCCGCCCA 1425
Qy 391 TGTGCTCAGAACTCAGAGTGGCTGCTTTAGA--TCACACATGCTCAAGTGTCATT 447
Db 1426 TGTGCTACAGAGCTCAACCGCAGTGAAGTCAAGTTCCTTACAGTTCACATGACCT 1485
Qy 448 ACCGAAATTCGAGATGTCAGAAAGCTGCTTACTGAACTTAATGGCTGCTGCCACG 507
Db 1486 TCCCTTCGACCTCAATGCTGAGACTGCTGTACTCAGCACTCAACGCCGCCCAATA 1545
Qy 508 CGATACAGGCTCCAAAGCTCCCGCTTGTGTGCTCCAGAGAGTGTGCTCAAGACTG 567
Db 1546 GCATGAAGGCTTAACTGCGGATGTGTGATGATCAGAGTGTGTGCTGTTTGTG 1605
Qy 568 GCTCAGCTCCATCTTGAATGGTCCGCTGCTGCTGATGAGAGAGCTGTGTTGTG 627
Db 1606 GATGGCTTCCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
Qy 628 TGTGCAATGATCGGCTAGGAATATTGTTCTTACCAACATGAGATCAGATGCTCGG 687
Db 1666 TCATCAGATACCGCTGAGGTGTGCTGAGGCTTCTTACAGACTGAGACCAAGCAACG 1725
Qy 688 GGAATCGGGCTTCAAGAGCAAGGAGTGTGCTGCTGGGCTCCAGAAACATGAGT 747
Db 1726 GGAATCGGGCTTCAAGAGCAAGGAGTGTGCTGCTGGGCTCCAGAAATATGCGCC 1785
Qy 748 TCTTCGTGGGAGCCCAAGCTGTGACCATCTTGGGAGAGTCCGCGGAGGCATTAAGT 807
Db 1786 ACTTTGAGAGGCAACCTGACGCTGACCATATTTTGGGAGTCTCGGGTGGCAGAGTG 1845
Qy 808 TTTCTAGTCTTAATCTGCTCCATGAGCCAAAGCTTATTCACAAAGCATCATGAGAGA 867
Db 1846 TGTCTTCCTGTTGTGTGTCCTCCCATATCCCAAGGACTCTTCCACGAGCATCATGAGAGA 1905
Qy 868 GGGGGGTGGCATATCTCCCTTAACCTGAGAGGCCCATGATTATGAGAAAGTGAAGCTG 927
Db 1906 GTGGGTGGCTCTCTGCGCGGCTCATATGCTCAGCTCAGCTA-----TGTATCTTCCA 1959
Qy 928 AGTGTGTCACATTTCTGTGTGAACAATGCGTCAAGCTGAGAGCCCTGCTGAGTGGC 987
Db 1960 CGGTGTGGCCCAACTGTCTGCTGTGACCAAGTGAATCTGAGGCTCGGTGGCTGCG 2019
Qy 988 TGAGACAAACCTTCAAGAGAGTGTGACCTTGAAGCAAGAAACAAAGTCTTCACTC 1047
Db 2020 TCGGGGGCAAGATGAAGAGGATTTGCAATTAACAAAGCTTCAAGATGATCCCG 2079

Db 1714 TCATTCAGTACCGCTGGGTGTCTGGGCTTCTTCAGCACTGGAGCAAGCAAGCAACCG 1773

Qy 688 GGAATCTGGGCTTCAAGGACGAGTGGCTCTGTCCCTGGGTCCAGAAAGCATCGAGT 747

Db 1774 GCAATCTGGGCTTCAAGGACGAGTGGCTCTGTCCCTGGGTCCAGAAAGCATCGAGT 1833

Qy 748 TCTTCGATGGGACCCGAGCTCTGTGACATCTTGGGAGAGTCCGGGAGCCATAGT 807

Db 1834 ACTTTGAGGCAACCTGACCGGTCTACCATTTTGGGAGTCTGGGGTGGCAAGAGT 1893

Qy 808 TTTCTAGTCTTATCTGTCTCTCCATGGCCAAAGCTTATTCACAAAGCCATATGAGA 867

Db 1894 TGTCTCTCTGTGTGTGTCTCCCATATCCCAAGGACTCTTCCAGGACCATATGAGA 1953

Qy 868 GTGGGGTGGCATATCTCCCTTACCTGGAGGCCATGATTATGAGAAGTGAAGCTGCG 927

Db 1954 GTGGGGTGGCATATCTCCCTTACCTGGAGGCCATGATTATGAGAAGTGAAGCTGCG 2007

Qy 928 AGGTGGTGGCAATTTCTGGGTAACAATGCGTCAAGCTCTGAGGCGCTGAGAGTGC 987

Db 2008 CGGTGGTGGCAACCTGTCTGCTGTGACCAAGTGAATCTGAGGCGCTGAGAGTGC 2067

Qy 988 TGAGGACAAACCTCTCAAGAGCTGTGACCTCAGCCAGAAACAAAGTCTTCACTC 1047

Db 2068 TGCGGGGCAAGATTAAGAGATTTCTTGCAATTAAGGCTTCAAGATGATCCCG 2127

Qy 1048 GAGTGGTGAATGAGTCTTCTTCTTCTTATGAGCTCTGATCTATTTGTCTAGAAAGAT 1107

Db 2128 GAGTGGTGAATGAGTCTTCTTCTTCTTATGAGCTCTGATCTATTTGTCTAGAAAGAT 2187

Qy 1108 TTAAGCAATTTCTTCCATCATCGGAGTCAATTAACAGAGTGTGCTTCTGCTGCC-- 1165

Db 2188 TTTCAAGCTGTCTCTAGCAATGTGTGTCTCAACAAGATTAAGATTCGCTGCTATCCCA 2247

Qy 1166 -----TATGAGAGAGGCTCTGAGATCTCAGTGGCTCAACAGATCCCTTGCCC 1215

Db 2248 AGGTATGAGATCTATGATATCCAGAAAGAAATGAGAGAGGCTCCAGGCTGCTC 2307

Qy 1216 TCCATCTGATCAAAACATCTCTGACATCCGCGCTCAGTATTTGACCTTGTGGTAAATG 1275

Db 2308 TGCAGAAATGTTAAAGCTGTGATGTGTGCTCTCAACATCTGTGATCCTGTGAGGAGG 2367

Qy 1276 AATATCTTCATGACAAAGCACTCCCTGACAGTAAATCCGAGACAGTCTTGTGACTTGTG 1335

Db 2368 AGTACATTTGGGACAAATGGGATCCCAAGACCTCTCAAGCGAGTTCAGAGAGATGATGG 2427

Qy 1336 GAGATGTGTCTTGTGTGTCTCTGACATGATCAAGCTGATATCAAGAGATGCTGTG 1395

Db 2428 CGGACTCATGTTGTGTGATCCCTGACATCCAGTAGCAATTTTCAGTG---TTCGCGGG 2484

Qy 1396 CACTGTCTTACTTATGATGATTTGGGACCGGCTCAAGTCTTGTGAAGACAGAGCCGG 1455

Db 2485 CCCCCTGTACTTCTTCAAGATTCAGATCCAGCTGTGCTCAAGAAACATCAAGCCAC 2544

Qy 1456 CTTTGTCAAGCCGACCAAGCTGATGAAGTCCGTTTGTGTGTGTGTGCTTCTGTA 1515

Db 2545 CGCATATTAAGCCAGACATGTTAA----- 2569

Qy 1516 AGGGGACATTTATTTTCAGAGAGCCAGAGAGAGAGATTAAGTCAAGCCGAGAA 1575

Db 2570 -----ATTCACTGAGAGAGAGAGAGAGATTAAGCAGAGAA 2604

Qy 1576 TGAATGAATATCTGGGCTACTTTGTCTGAAACCGGGAATCTTAATGGAAAGCACTGTCTC 1635

Db 2605 TGAATGAATATCTGGGCTACTTTGTCTGAAATGGGAATCCCAATGGCAGAGGTCTGCAAC 2664

Qy 1636 TGTGGGACGTTATTAATCTGATGAGAGTACCTCAAGTGAATCTTGAACATGAGCTCG 1695

Db 2665 ACTGGCCGCTGTTCAGACAGAGAGAGCAATACCTCGGCTGAACCTTAAGCTGTGGGTGG 2724

Qy 1696 GACAGAGACTCAAGAAACCGCGGGTGAATTTTGGACAGCAACATCCGCC 1746

Db 2725 GCGGGGCTCTGAAGGCCCAAGGCTCCAGTTCTTGAAGAGAGCGTGCGCC 2775

RESULT 13

CE111083

LOCUS

DEFINITION

Shultzomic04334 Rat lung airway and parenchyma cDNA libraries

Rattus norvegicus cDNA clone Contig3845 5', mRNA sequence.

CE111083

VERSION

CE111083.1 GI:33167595

KEYWORDS

EST.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 1918)

Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V., Padua,A.M., Gurske,W.A., Morin,D., Penn,S.G., Jovanovich,S.B., Plopper,C.G. and Buckpitt,A.R.

Gene expression analysis in response to lung toxicants: I. Sequencing and microarray development

Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)

CONTACT

Dr. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)

Dept. of Molecular Biochemistry, School of Veterinary Medicine University of California, Davis

1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA

Tel: 530 752 0793

Fax: 530 752 4698

Email: mshultz@ucdavis.edu

Average Phred score is 20 or better. All poor quality data (Phred < 20) and vector/linker sequence has been removed.

High quality sequence atop: 1918.

FEATURES

source

1..1918

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="Contig3845"

/sex="male"

/tissue_type="airway or parenchyma"

/dev_stage="adult"

/clone_lib="Rat lung airway and parenchyma cDNA libraries"

/note="Organ: Lung; Vector: pGEM-11zf(-); Site 1: Eco RI; Site 2: Not I; mRNA was isolated from microdissected rat lung airways and parenchyma tissues."

ORIGIN

Query Match 15.7%; Score 338; DB 7; Length 1918;

Best Local Similarity 54.2%; Pred. No. 3.4e-80;

Matches 869; Conservative 0; Mismatches 660; Indels 75; Gaps 6;

Qy 186 GGGCTTGTGTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGCAAGCA 245

Db 83 GGGATATCATCTCTCAACACCTGTGTGAACAATGTAAGGCAAGTCTGGGAAAGTAT 142

Qy 246 GTCACTGTCTGGGAAGCCCTGTGTGTGAACGTGTCTCTGAGTCCCTTGTCTGCT 305

Db 143 GTCAATTTGGAAGGATTTGACAGCTGTGTGTGTCTGTGGAATCCCTTGCACAG 202

Qy 306 CCCCCGTGGATCCCTGATTAAGAACCCGAGCTGTGATGCGCTGGGATTAATCTTG 365

Db 203 TCCCTCTTGGGCTCTTGAAGTTGCTCTCAACACAGCTGTGAGATCTTGAATCTTG 262

Qy 366 CGAAGGACACTCTTACCTTAATTTGTGCTCCAGAACTC-----AGATGCTGTCTC 419

Db 263 AAGATATTAATCTCTTACCTTACCTATGCTCTCAAGATGCTGTGAGGCAAGTCTCTC 322

Qy 420 TTAATCAACAGATCTTAAGTCAATTAACCGAAATTCGAGTCTCAAGAGCTGCTC 479

Db 323 TCAGAGCTTTTACCAACAGAGAGAAACATTCCTTTCAGATTTTTCAGAGATGCTCTC 382

Qy 480 TACTGAACATCTATGCGCTGCCAGCGATACAGGCTCCAAGCTCCCGTCTTGTG 539

Db	383	TACTGTAATGTTTATCTCCGCTGACTTGAACAAGAACAGCCGGTTACAGAGATGCTG	442
Oy	540	TGGTTCACAGAGGTGCTTCAGAGCTGGCTCAGCTTCATCTTTGATGGGTCGCCCTG	599
Db	443	TGGATCATGAGGTGAGTGAATGAGTGGATGATCATCTATGATGAGACAGGCCCTC	502
Oy	600	GCTGGCTATGAGAGAGTGTGTTGGTTCCTCCAGTACGGGCTAGAGAAATTTGGTTTC	659
Db	503	TCTGGCCATGAAAAGTGGTGGTGGTGAACATTCAGTATGCTCTGGCATCTGGGGATTC	562
Oy	660	TTACACCAATGGGATCAGATGCTCCGGGGAAGTGGGCTTCACAGACCAAGTGGCTGCT	719
Db	563	TTACAGACAGGGGATGAAACACAGCCAGGGCAACTGGGGTCAATTGGACACAGCTGGCTG	622
Oy	720	CTGTCTGGGTTCAGAAAACATCGAGTTCTTCGGTGGGGAACCCAGCTCTGTGACATC	779
Db	623	CTGACCTGGGTTCAGAGAAATATGTCACAACTTTGGGGGCAACCCAGGCTCTGACATC	682
Oy	780	TTTGGCGAGTCCGCGGGAGCCATTAAGTGTCTTCTATGCTATACGTCTCCCATGGCCAAA	839
Db	683	TTTGGAGATCTGACAGAGGTTTCAAGTCTCTCTTTCTTGTATATCTCTCTGGGCCAAG	742
Oy	840	GCGTTATTTCCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGGCC	899
Db	743	AACCTCTTCCACAGGGCCATTTCTGAGAGTGGTGTCTCCCTCAATCTGCTGATTAACA	802
Oy	900	CATGATTTAGAGAGAGTGAAGACCTCAGGTGGTGGCAATTTCTGTGTGAACATGCG	959
Db	803	ACAGATGACAAAGCCATTTCTATCTGATTTGCTACTTTTCTGGGTGTAACACACACA	862
Oy	960	TCAGACTCTGAGGCCCTGCTGAGGTGGCTGAGACA-----	995
Db	863	TCAGGTGTTATGTTCAATGCTGCGCCAGAAAGACAGAGATGAATCTCGGAGACTTCA	922
Oy	996	-----AAACCTCCAGAGAGCTGTGACCTCAGCCAGACCAAAACAAGTCT-----	1040
Db	923	TTTAAATTTGAATCTTTTAACTGAGACCTATTTTGAACCCAAAGAGCTATCCCTTC	982
Oy	1041	TTCACTGAGTGTGTTAGTGTCTTTCTTCTTAATGAGCCTTAGATCTATGTTCTGAG	1100
Db	983	CTACCTACTGTGATTTAGTGAAGTGTGTCTCTCAAAAGACACAGAAAGATCTGGCTGAG	1042
Oy	1101	AAAGATTTAAGCAATTTCTTCATCATGATGAGATCAATPACACGAGTGGCTTCTCTG	1160
Db	1043	AAGAGTTTCAACAGCGTCCCTCATGTGTGGGATCAACAGACAAAGTTTGGCTGATC	1102
Oy	1161	CTGGCTATG-----AAGAGGCTCTGAGATCTCAGTGGTCCAAACAATTCCT	1210
Db	1103	ATTTCATGCTTATGGGCTATTCACCTCTCCGAAAGCAAACTGAGCACAAAACACCCAT	1162
Oy	1211	TGCGCTCATCTGATACAAAC--ATCTGACATCCGCGCTCAGTATTTGACCTTGTG	1268
Db	1163	TCCCTCTTGTGAAAGTCTTACCCACACATGAAATATCTGAGAAATATGATTCATAGCT	1222
Oy	1269	GCTATATGATTACTTCCATGACAGACCTCCCTGACATGAAATCCGAGACAGTCTTGTGAC	1328
Db	1223	TCTGAGAAATTTAGAGGGAACAGATGACCTCGGAAAGAAAGACCTGTTCAAAGAC	1282
Oy	1329	TTGCTTGGAGATGTGTTCTTTGTGTCCCTGACCTGATCAAGCTCATACAGAGAT	1388
Db	1283	TTGGTTGCAATGTGATATTTGGTGTCCCATGATGATGTGTCTCGAAGTACAGAGAT	1342
Oy	1389	GCTGGTGCACCTGTCTATCTTATGAGTTTGGCACCGGCTCAGTGGCTTTGAGACAG	1448
Db	1343	GCTGAGCCCCCACCTTCATGATGATTTGAGTATCGCCAAAGCTTTGTATCAGCCATG	1402
Oy	1449	AAGCGGCTTTTGTCAAAGCCAGCAGCTGATGAAGTCCGCTTTGTGTGGTGGTGGCC	1508
Db	1403	AGGCCCAAGACAGTGAATCGAGAGCAGATGATGAAATTTCTTCAATTTTGGATCTCCA	1462
Oy	1509	TTCTCTGAAGGGGACATTTGTATGTTGAAAGGCCACGAGAGAGAGAAAGTTACTGAGC	1568

Db	1463	TTTTTAA-----AGATGGCGGCTTCAGAGAGAGAACCATCTCAGC	1504		
Oy	1569	CGAAGATGATGAATACTGGGCTACTTGTGTCGAAACCGGGAATCTTAATGGGAACGAC	1628		
Db	1505	AAAATGATGATGAATCTGGGCCAATTGTCTCGGAATGGGAACCTTAATGGGGAGGCG	1564		
Oy	1629	CTGTCTGTGGCCAGCTTATATCTGACTGAGACGATACCTTCAGCTGACCTTGAACATG	1688		
Db	1565	CTGGCCCATTTGGCCAGAAATATGACAGAAAGAGGATCTGAAGATTTGGTCTCAACT	1624		
Oy	1689	AGCCTGGACAGAGACTCAAAAGAACCGCGGGGATTTTGTGAC	1732		
Db	1625	CAGCAGCCCAAGGCTTGAAGCAAAAGATGTCTTTTGTGTC	1668		
Db	AK007235	989 bp	mRNA	linear	HTC 03-Apr-2004
Db	AK007235		Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700122C07 product:serine/threonine kinase 33, full insert sequence.		
Db	AK007235		AK007235		
Db	AK007235.1	GI:12840656	HTC; CAP trapper.		
Db	AK007235		Mus musculus (house mouse)		
Db	AK007235		Mus musculus		
Db	AK007235		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
Db	AK007235		1		
Db	AK007235		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
Db	AK007235		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
Db	AK007235		Genome Res. 10 (10), 1617-1630 (2000)		
Db	AK007235		20499374		
Db	AK007235		11042159		
Db	AK007235		3		
Db	AK007235		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komo, H., Akiyama, J., Nishi, K., Kikunishi, T., Taepilo, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.		
Db	AK007235		RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer		
Db	AK007235		Genome Res. 10 (11), 1757-1771 (2000)		
Db	AK007235		20530913		
Db	AK007235		11076861		
Db	AK007235		4		
Db	AK007235		The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
Db	AK007235		Functional annotation of a full-length mouse cDNA collection		
Db	AK007235		Nature 409, 685-690 (2001)		
Db	AK007235		5		
Db	AK007235		The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
Db	AK007235		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
Db	AK007235		Nature 420, 563-573 (2002)		
Db	AK007235		6 (bases 1 to 989)		
Db	AK007235		Adechi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komo, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, Y.,		

Db	1463	TTTTTAA-----	-AGATGGCCCTCAGAAAGAGACCAATCTCAGC	1504
Oy	1569	CGGAAGATGATGAATACTGGGCTACCTTTGCTCGAACCGGGATTCCTAATGGGAAGAC		1628
Db	1505	AAATGTGATGAATAATCTGGGCCAATTGCTCTCGAATGGGAACCTTAATGGGGAGGC		1564
Oy	1629	CTGTCTCTGTGGCCAGCTTATTAATCTGACTGAGACGTACTCCAGCTTGACCTTGACATG		1688
Db	1565	CTGCCCATTTGGCCAGATATGACCAAGAAAGAGGGTACTGAAAGATTGGTGCTCACT		1624
Oy	1689	AGCCTCGACAGACGACTCAAGAAACCGCGGTGATTTTGGAC	1732	
Db	1625	CAGCAGCCCAAGAGCTGAAGACAAGAAAGTGGCTTTTGTGTC	1668	
RESULT 14				
AKO07235				
LOCUS				
DEFINITION		989 bp	mRNA	linear
				HTC 03-APR-2004
				Mus musculus adult male testis cDNA, RIKEN full-length enriched
				library, clone:1700122C07 product:serine/threonine kinase 33, full
				insert sequence.
ACCESSION				
VERSION	AKO07235.1	GI:12840656		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE				
AUTHORS	1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Methods. Enzymol. 303, 19-44 (1999)			
MEDLINE	98279253			
PubMed	10349636			
REFERENCE				
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to			
JOURNAL	prepare full-length cDNA libraries for rapid discovery of new genes			
MEDLINE	20499374			
PubMed	11042159			
REFERENCE				
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
	Kono, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,			
	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,			
	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			
	Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,			
	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,			
	Okazaki, Y., Murakatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format			
JOURNAL	sequencing pipeline with 384 multicapillary sequencer			
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)			
PubMed	20530913			
REFERENCE				
AUTHORS	4 11076861			
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the			
JOURNAL	PANTOM Consortium.			
REFERENCE	Functional annotation of a full-length mouse cDNA collection			
AUTHORS	Nature 409, 665-690 (2001)			
TITLE	5 The PANTOM Consortium and the RIKEN Genome Exploration Research			
JOURNAL	Group Phase I & II Team.			
REFERENCE	Analysis of the mouse transcriptome based on functional annotation			
AUTHORS	of 60,770 full-length cDNAs			
	Nature 420, 563-573 (2002)			
	6 (bases 1 to 989)			
	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,			
	Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,			
	Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,			
	Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,			
	Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,			
	Koyu, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, Y.			


```

/mol_type="mrna"
/db_xref="FVB/N"
/clone="IMAG:5123923"
/tissue="Liver, normal, 5 month old male mouse."
/clone_id="NCI CGAP L19"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"

```

```

Query Match      15.4%; Score 331.4; DB 3; Length 2038;
Best Local Similarity 53.1%; Pred. No. 2.2e-78;
Matches 839; Conservative 0; Mismatches 711; Indels 31; Gaps 5;

```

```

QY 183 ACTGGGCTTCTGCTGAAGGGCCACAGAGAACCCAGGCTGGAGTGGATTGACGGCAAG 242
DB 98 ACTGGACCCAAAGTCAATTCAGCTGAGTGGATTACCCCTGGGCTGTGTTCAAGGCCG 157
QY 243 CAAGTCACTGTCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCGAGTCCCTTTGCT 302
DB 158 CAGGTGGGTGTGAAGGACAGACCCGATGTAAATGTCTTCTGGGCAATCCCTTTGCT 217
QY 303 GCTCCCGGCTGGAGTCCCTGCAATTTAGCAACCCGACCTGCAATCCCTGGATTAAC 362
DB 218 CAAGCACAGTGGAACTCTTCGTTCTAGCTTCACTCCACACAGAGCCCTGGAGAGT 277
QY 363 TTGCGAAGAACCACTCTTAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTA 422
DB 278 GTGAGAGATGCCAGATCAATCCCAATGTGCTTCAAGATGTAGAAATGATCAAC 337
QY 423 GATCAACATGTGTCAAGTGCATTACCCGAATTTGAGTGTCAAGAGATGCTCTTAC 482
DB 338 AGCAGATTTGGCTTCATATAAAGATTAAATCTTCCCATTTCTGAGAGATGCTCTGACC 397
QY 483 CTGAACATCTATGCGCTGCTCCACCGCATACAGGCTCCAGCTCCCGTCTTGGATGG 542
DB 398 CTCAACATCTACAGGCCCACTGAGATCACTGACAGGGAACAAAGCCGCTCATGTATGG 457
QY 543 TTCCCAAGAGTGGCTCTTCAAGACTGCTCAAGCTTCATTTGATGGTCCGCTGCT 602
DB 458 ATCCACGAGAGGCTCTCTGTGTGGTCTCTCCATCATCTCAGAGATGATCAGCACTGGCC 517
QY 603 GCTTATGAGAGCGTGTGTTGTGTCGTCAGTACCGGCTAGAGAAATTTGTGTTCTTC 662
DB 518 GGTATCGGGAGATGTGTATGTCTACTGTCCAGTATCCGCTTGGGATCTTTGGCTCTC 577
QY 663 ACCACATGGGATCAGCATGCTCCGAGGAACTGGGCTTCAAGGACAGAGTGGTCTG 722
DB 578 AGCATGTGAGACAGACATGCTCAGGCAACAGGGGATTCGTGATGTGGTGTCTT 637
QY 723 TCTTGGGTCCAGAAAGATGAGTCTTTCGATGGG-GACCCAGCTCTGTGACCATCTT 781
DB 638 CGCTGGGTCCAGGGGAACATAGCCCCCTTTGGGGGTATCCCACTGTGTCACTATCTT 697
QY 782 TGGCGAGTCCGCGGAGGCATATAGTGTCTTATCTGTCTCCATGAGCCAAAG 841
DB 698 CGGTAACTCTGTGAGGATGATGTCTATCCCTGTCTGTCTCAATATGTGCTGG 757
QY 842 CTATTTCCAAAGCCATCATGAGAGTGGGAGCCATATCCCTTACCTGAGGCCCA 901
DB 758 GCTCTTCCAGAGACCATATCCAGAGTGGATTTGTCAAGCATATATGATGG-----A 811
QY 902 TGATTATGAGAAAGTGAAGACTGCAAGTGTGCAATTTCTGTGTAAACAATGCTC 961
DB 812 AGACATGAACCAATGGCTGTGAAGCTCAGAACTTTGCAATTTCTGTGGCTGTGTGC 871
QY 962 AGACTCTGAGGCTCTGTGAGTGTCTGAGACAAAACCTTCAAGAGAGCTGTGACCT 1021
DB 872 ATCCCGAGCTGAGTGTGTGCTGTGCTGAGAAAGAAAGGACCTT-----AT 925
QY 1022 CAGCCAGAAAACAAAGTCTTCACTGAGTGTGATGTGCTTTCTTTCTATATAGCC 1081
DB 926 CAGCGAAGAAAAGTGAACATTTCTTACATATGATGATGCTCTTCTTCCACAAAGGCC 985

```

```

QY 1082 TCTAGATCTATGTCTCAGAAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATA 1141
DB 986 AAGAAAGCTCCATAGAGACACGACATTTCCACATGTGCTTCACTTTTGGAGTCAACAA 1045
QY 1142 CCAGAGTGTGGCTTCTGCTG-----TGCCTATGAAGAGGCTCTGAGATCTT 1189
DB 1046 CCATGAATTTGGCTGGCTTCTTCACTCAATTCCTTGAATTCCTGATTAAGTTGAACATCT 1105
QY 1190 CAGTGGCTCAACAAAGTCCCTTGGCCCTCATCTGATTAACAAAACATCTGACATCCGCC 1249
DB 1106 GAGCGGGGAAGACCTGTGGAGATTTCAAGGCCCTTCTTGGCCATTAAGAAAGGCCCC 1165
QY 1250 TCAATTTTGAACCTTGTGGCTAAATGAATATCTTCATGACAAAGCACTCCCTGAAT 1309
DB 1166 TGAGATCATGCCCCATGTCATGATGAATATCTAGACAAATGGCTCAATCATATAGCTAC 1225
QY 1310 CCGAGACAGTCTTCTGACATTTGCTTGGAGATGTCTTTGTGTGCTCCCTGCACTGATCAC 1369
DB 1226 AAGGTATGCTTCCAGAAATTTGTGGGTGATATCTGTTATCATATCTTACCTTGAACCTT 1285
QY 1370 AGCTGATATCAGAGATGTGGTGCACCTGTCTACTTCTATGAGTTTGGACACGGGCC 1429
DB 1286 CTCAAAATACCTTGAAGATGTGGGTGCTGCTGTTTCTTGTAGAGTTTCAAGCATACACC 1345
QY 1430 TCAATGCTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCG 1489
DB 1346 CAGTCTTTTGAAGATTCAAGCCAGCTGGGTGAAGGCTGACCAATGCGTCTGAGAAATT 1405
QY 1490 CTTTGTGTGGTGGTGGCTTCTCTGA-----AGGGGGAATGTGTTATGTTGCAAGAGC 1543
DB 1406 CTTTGTTTGTGAGGTCTTCTTCTCACTGATGAGATGTTCCCTTGTGGCTTCCCAAGGC 1465
QY 1544 CACGAGAGAGAAAGTTACTGAGCCGGAAGATGAATACTGGCTTACCTTTGCTG 1603
DB 1466 CACAGAGAAAGAAAGAGAGCTGAGCTGACATGATGAGCCCAATGAGCAAGTTGCAAG 1525
QY 1604 AACCGGGAATCTTAATGGGAAGAGCTGTCTGTGGGCCAGCTTATATCTGACTGAGCA 1663
DB 1526 CACAGAAACCCCAATGGCAAGGGGCTGCTCTTGTGGCCCAATTAACCAAGTTAGMAA 1585
QY 1664 GTACTTCCAGCTGAACTTGAACATGAGCCTCGACAGAGACTCAAGAAACCGCGGGTGA 1723
DB 1586 ATACTTGAAGATGTGTGAATCAACGAGCTGGGAGTGAAGCTTAAGAGGATCGCTACA 1645
QY 1724 TTTTGGACAGACCATCTCC 1744
DB 1646 GTTCTGAGAGAGACCTGCC 1666

```

Search completed: June 15, 2005, 02:39:37
 Job time : 7174.54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 359.298 Seconds

(without alignment)
9827.730 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 2158
Sequence: 1 ccacgcgtccgcaaaacagc.....aaaaaaaaaaaaaaaaaaaaa 2158

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:
1: /cgn2_6/pdata/1/ina/5B_COMB.seq:*
2: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
5: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
6: /cgn2_6/pdata/1/ina/6B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	100.0	2158	US-10-023-515-1	Sequence 1, Appl1
2	1746	80.9	1746	US-10-023-515-3	Sequence 3, Appl1
3	508.6	23.6	965	US-09-799-451-155	Sequence 155, App
4	406.6	18.8	2117	US-09-949-016-3799	Sequence 3799, Ap
5	406.6	18.8	2169	US-09-949-016-555	Sequence 555, App
6	406.6	18.8	2191	US-09-595-6828-27	Sequence 27, Appl
7	362.6	16.8	1717	US-09-595-6828-20	Sequence 20, Appl
8	357	16.5	1701	US-09-264-737-3	Sequence 3, Appl1
9	176	8.2	16897	US-09-949-016-13807	Sequence 13807, A
10	139.6	6.5	5767	US-09-810-861B-3	Sequence 3, Appl1
11	139.6	6.5	14446	US-09-810-861B-4	Sequence 4, Appl1
12	139.4	6.5	1725	US-09-810-861B-5	Sequence 5, Appl1
13	139.4	6.5	1845	US-07-732-962A-1	Sequence 1, Appl1
14	139.4	6.5	1845	PCT-US92-06106-1	Sequence 1, Appl1
15	139.4	6.5	2158	US-09-949-016-1192	Sequence 1192, Ap
16	139.4	6.5	2158	US-09-949-016-1193	Sequence 1193, Ap
17	139.4	6.5	2256	US-08-318-826A-5	Sequence 5, Appl1
18	139.4	6.5	2256	US-08-318-826A-5	Sequence 5, Appl1
19	139.4	6.5	2256	US-08-318-826A-5	Sequence 5, Appl1
20	139.4	6.5	3016	US-08-318-826A-7	Sequence 7, Appl1
21	139.4	6.5	3016	US-08-370-156-5	Sequence 5, Appl1
22	139.4	6.5	3016	US-08-370-156-5	Sequence 5, Appl1
23	139.4	6.5	3096	US-08-318-826A-6	Sequence 6, Appl1
24	139.4	6.5	3096	US-08-370-156-3	Sequence 3, Appl1
25	139.4	6.5	3096	US-08-814-095-3	Sequence 3, Appl1
26	139.4	6.5	9885	US-09-949-016-12934	Sequence 12934, A
27	139.4	6.5	9885	US-09-949-016-12935	Sequence 12935, A

28	139.4	6.5	35060	US-08-814-095-7	Sequence 7, Appl1
29	128.6	6.0	12827	US-09-949-016-15541	Sequence 15541, A
30	123	5.7	1764	US-09-491-356C-3	Sequence 3, Appl1
31	123	5.7	2802	US-09-949-016-2515	Sequence 2515, Ap
32	123	5.7	2862	US-09-949-016-2516	Sequence 2516, Ap
33	123	5.7	4436	US-09-491-356C-6	Sequence 6, Appl1
34	121.4	5.6	4233	US-09-491-356C-5	Sequence 5, Appl1
35	105.6	4.9	836	US-09-270-767-861	Sequence 861, App
36	105.6	4.9	836	US-09-270-767-16143	Sequence 16143, A
37	101.2	4.7	1905	US-09-347-878-33	Sequence 33, Appl
38	101.2	4.7	1907	US-08-462-884A-2	Sequence 2, Appl1
39	101.2	4.7	1908	US-08-461-881B-2	Sequence 2, Appl1
40	101.2	4.7	1908	US-09-123-960-2	Sequence 2, Appl1
41	101.2	4.7	2184	US-08-445-050-8	Sequence 8, Appl1
42	101.2	4.7	2184	US-08-204-691-8	Sequence 8, Appl1
43	101.2	4.7	2375	US-09-949-016-3976	Sequence 3976, Ap
44	101.2	4.7	2428	US-08-445-050-1	Sequence 1, Appl1
45	101.2	4.7	2428	US-08-204-691-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-10-023-515-1
; Sequence 1, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 53010. A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-023-515-1
Query Match 100.0%; Score 2158; DB 4; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCACGCGTCCGAAAACAGCGCTGGAGAGCATGTGAGTAAAGTAAATTAACGA 60
1 CCACGCGTCCGAAAACAGCGCTGGAGAGCATGTGAGTAAATTAACGA 60
61 TTTAAATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
61 TTTAAATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
121 CTTCAATGCTGCTTTTCTGATTTCTGATTTCTGATTTCTGATTTCTGATTT 180
121 CTTCAATGCTGCTTTTCTGATTTCTGATTTCTGATTTCTGATTTCTGATTT 180
181 AAACGTGGGCTTCTGATTTCTGATTTCTGATTTCTGATTTCTGATTTCTGAT 240
181 AAACGTGGGCTTCTGATTTCTGATTTCTGATTTCTGATTTCTGATTTCTGAT 240
241 AGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCC 300
241 AGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCC 300
```

```
QY 301 CTGCTCCCCGCTGGAGATCCCTGCGATTTACGAACCCGAGCTGATCGCCCTGGAGATA 360
Db 301 CTGCTCCCCGCTGGAGATCCCTGCGATTTACGAACCCGAGCTGATCGCCCTGGAGATA 360
QY 361 ACTTGGAGAGAGCCACCTCTACCCCTAATTTGTGCTCTCAGAACTCAGAGTGGCTCTCT 420
Db 361 ACTTGGAGAGAGCCACCTCTACCCCTAATTTGTGCTCTCAGAACTCAGAGTGGCTCTCT 420
QY 421 TAGATCAACAGATGCTCAGAGTGCATTTACCCGAATTTGGAGTGTCAAGAGATGCTCTCT 480
Db 421 TAGATCAACAGATGCTCAGAGTGCATTTACCCGAATTTGGAGTGTCAAGAGATGCTCTCT 480
QY 481 ACCTGAACATCTATGCGCTGCGCCAGAGCTCAAGCTCCAGAGTGTCCCGCTTGTGTGT 540
Db 481 ACCTGAACATCTATGCGCTGCGCCAGAGCTCAAGCTCCAGAGTGTCCCGCTTGTGTGT 540
QY 541 GGTTCGCGAGAGTGGCTCTTCAAGAGCTGCTCAAGCTCCATCTTTGATGGTCCGCTGTG 600
Db 541 GGTTCGCGAGAGTGGCTCTTCAAGAGCTGCTCAAGCTCCATCTTTGATGGTCCGCTGTG 600
QY 601 CTGCTATGAGAGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 601 CTGCTATGAGAGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 661 TCACCACATGGGATCAGATGCTCCGCGGAACTGGGCTTCAAGAGACAGAGTGGCTGTCT 720
Db 661 TCACCACATGGGATCAGATGCTCCGCGGAACTGGGCTTCAAGAGACAGAGTGGCTGTCT 720
QY 721 TGTCTGGGTTCAGAGAAATCGAGTCTTCCGTGGGAGCCCACTGCTGTGACATCTT 780
Db 721 TGTCTGGGTTCAGAGAAATCGAGTCTTCCGTGGGAGCCCACTGCTGTGACATCTT 780
QY 781 TTGGCGAGTCCGCGGAGCCATAGTGTCTTAGTCTTATAGTCTTCCATGCGCAAG 840
Db 781 TTGGCGAGTCCGCGGAGCCATAGTGTCTTAGTCTTATAGTCTTCCATGCGCAAG 840
QY 841 GCTTATTCACAAGCATCATGAGAGAGTGGGCTGCTCATCTCTTCTGAGAGCC 900
Db 841 GCTTATTCACAAGCATCATGAGAGAGTGGGCTGCTCATCTCTTCTGAGAGCC 900
QY 901 ATGATTTATGAAAGATGAGAGCTTGCAGGTGTGCAATTTCTGTGTGTAACAATGGGT 960
Db 901 ATGATTTATGAAAGATGAGAGCTTGCAGGTGTGCAATTTCTGTGTGTAACAATGGGT 960
QY 961 CAGACTCTGAGGCGCTGCTGAGAGTGTGCTGAGAGCAAAACCTTCCAGAGAGCTGTGACC 1020
Db 961 CAGACTCTGAGGCGCTGCTGAGAGTGTGCTGAGAGCAAAACCTTCCAGAGAGCTGTGACC 1020
QY 1021 TCAGCCAGAAACAAAGTCTTCACTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 TCAGCCAGAAACAAAGTCTTCACTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1081 CTCTAGATCTATGTCTGAGAAAGATTTAAAGCAATCTTCCATCATCGAGTCAATA 1140
Db 1081 CTCTAGATCTATGTCTGAGAAAGATTTAAAGCAATCTTCCATCATCGAGTCAATA 1140
QY 1141 ACCACGAGTGTGCTTCTGCTGCTATGAGAGAGCTCTGAGATCTCTAGTGGCTCCA 1200
Db 1141 ACCACGAGTGTGCTTCTGCTGCTATGAGAGAGCTCTGAGATCTCTAGTGGCTCCA 1200
QY 1201 ACAAGTCCCTGCGCTCATCTGATTAACAAAATCTGCAATCTCCGCTCAGATTTGG 1260
Db 1201 ACAAGTCCCTGCGCTCATCTGATTAACAAAATCTGCAATCTCCGCTCAGATTTGG 1260
QY 1261 ACCTGTGGCTAATGAATACTTCCATGACAGAGCTCCCTGAGTGAATCCGAGAGCTC 1320
Db 1261 ACCTGTGGCTAATGAATACTTCCATGACAGAGCTCCCTGAGTGAATCCGAGAGCTC 1320
QY 1321 TTCTGAGCTTGTGGAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 1321 TTCTGAGCTTGTGGAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
```

```
QY 1381 ACAGAGATGCTGTGCACTGTCTACTTATGAGTTCGAGCCGCTCAGTGTGTG 1440
Db 1381 ACAGAGATGCTGTGCACTGTCTACTTATGAGTTCGAGCCGCTCAGTGTGTG 1440
QY 1441 AAGACAGAGAGCGCTTTTGTCAAGCGAGCAGCTGATGAATCCGCTTTGTGTG 1500
Db 1441 AAGACAGAGAGCGCTTTTGTCAAGCGAGCAGCTGATGAATCCGCTTTGTGTG 1500
QY 1501 GTGGTCCCTTCTGAAAGGGGAGCATTTGATGTTGGAAGAGCCAGAGAGAGAA 1560
Db 1501 GTGGTCCCTTCTGAAAGGGGAGCATTTGATGTTGGAAGAGCCAGAGAGAGAA 1560
QY 1561 TACTGAGCGAGAGATGATGAATPACTGGGTACTTTGCTCGGAACCGGGAATCTTATG 1620
Db 1561 TACTGAGCGAGAGATGATGAATPACTGGGTACTTTGCTCGGAACCGGGAATCTTATG 1620
QY 1621 GGAAGACCTGTCTGTGTGCTGAGCTTAATCTGACTGAGCAGTACTCAGCTGACT 1680
Db 1621 GGAAGACCTGTCTGTGTGCTGAGCTTAATCTGACTGAGCAGTACTCAGCTGACT 1680
QY 1681 TGAACATGAGCTGTGAGCAGAGACTCAAGAAACCGCGGTGAGATTTTGGACAGACCA 1740
Db 1681 TGAACATGAGCTGTGAGCAGAGACTCAAGAAACCGCGGTGAGATTTTGGACAGACCA 1740
QY 1741 TCCCTGATCTGTGTGCTGCTCGACATGCTCCAGAGTCTCTTCTTAACTTTCC 1800
Db 1741 TCCCTGATCTGTGTGCTGCTCGACATGCTCCAGAGTCTCTTCTTAACTTTCC 1800
QY 1801 TCTCTCTCTCCAGCCCTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
Db 1801 TCTCTCTCTCCAGCCCTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
QY 1861 TTGGTTCCTCTCTCTCCATTAATTTCTCCGCAATATGATGCTTCTTGTAGCTCA 1920
Db 1861 TTGGTTCCTCTCTCTCCATTAATTTCTCCGCAATATGATGCTTCTTGTAGCTCA 1920
QY 1921 GCTGCTTCTATGAGAGATCTTGCAAAACAGCTGCTTCTGCTGATTTATGACTTA 1980
Db 1921 GCTGCTTCTATGAGAGATCTTGCAAAACAGCTGCTTCTGCTGATTTATGACTTA 1980
QY 1981 GGAATATCTTACAGAAATCTTTCACATCAAAAGTGCATTTGTCTGGAAGGGA 2040
Db 1981 GGAATATCTTACAGAAATCTTTCACATCAAAAGTGCATTTGTCTGGAAGGGA 2040
QY 2041 CAAGATTTCTTAAATTAATTTGGAAGAGGCTGCTTATGTTGCTAATAATGTTTT 2100
Db 2041 CAAGATTTCTTAAATTAATTTGGAAGAGGCTGCTTATGTTGCTAATAATGTTTT 2100
QY 2101 TGTAACTATATGAATTAATCAAGATGTAAATAGGAAAAAATTTTTTTTTT 2158
Db 2101 TGTAACTATATGAATTAATCAAGATGTAAATAGGAAAAAATTTTTTTTTT 2158
```

RESULT 2
US-10-023-515-3
; Sequence 3, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 5310, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-12201
; CURRENT APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1746

TYPE: DNA
ORGANISM: Homo sapiens
US-10-023-515-3

Query Match 80.9%; Score 1746; DB 4; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

96 ATGCCAGAGGAGCTTATCTTCAATCTGTTTCAATAGTGTGCTTTTCTGATTTCCGACGCC 155
1 ATGCCAGAGGAGCTTATCTTCAATCTGTTTCAATAGTGTGCTTTTCTGATTTCCGACGCC 60
156 CTGTTGGGACACAGACAGTGGGGAAAACTGGCCCTTCTGCTGAAGGGCCACAGAGAAC 215
61 CTGTTGGGACACAGACAGTGGGGAAAACTGGCCCTTCTGCTGAAGGGCCACAGAGAAC 120
216 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTG 275
121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTG 180
276 AACGTTCTCTGGAGTCCCTTTGCTGCTCCCGGCTGGAGTCCCTGGATTTAAGAAC 335
181 AACGTTCTCTGGAGTCCCTTTGCTGCTCCCGGCTGGAGTCCCTGGATTTAAGAAC 240
336 CCGCAGCTGCATCGCCCTGGGATTAATTGCGAAGACCACTCTAACCTTAATTGTGC 395
241 CCGCAGCTGCATCGCCCTGGGATTAATTGCGAAGACCACTCTAACCTTAATTGTGC 300
396 CTCGAGAACTCAGAGTGGCTCTTATGATCAACATGCTCAAGGTGATTAACCGAAA 455
301 CTCGAGAACTCAGAGTGGCTCTTATGATCAACATGCTCAAGGTGATTAACCGAAA 360
456 TTCGAGATCAGAGACAGTGGCTCTTACCTGACATCTATGCGCTCCGACCGCATACA 515
361 TTCGAGATCAGAGACAGTGGCTCTTACCTGACATCTATGCGCTCCGACCGCATACA 420
516 GGCTCCAAGCTCCCGCTCTTGGTGGTCCGAGAGTGGCTTCAAGACTGGCTCAGCC 575
421 GGCTCCAAGCTCCCGCTCTTGGTGGTCCGAGAGTGGCTTCAAGACTGGCTCAGCC 480
576 TCCATCTTTGATGGGTCCGCTGGCTGCTATGAGAGACGTGTGGTTGTGCTCAGC 635
481 TCCATCTTTGATGGGTCCGCTGGCTGCTATGAGAGACGTGTGGTTGTGCTCAGC 540
636 TACCGGCTAGAAATATTTGGTTCTTCAACAATGGGATCAGATGCTCCGGGAACTGG 695
541 TACCGGCTAGAAATATTTGGTTCTTCAACAATGGGATCAGATGCTCCGGGAACTGG 600
696 GCCTTCAAGGACAGGTGGCTGCTGCTGGGGTCCAGAGAACATCGAGTTCTTCGGGT 755
601 GCCTTCAAGGACAGGTGGCTGCTGCTGGGGTCCAGAGAACATCGAGTTCTTCGGGT 660
756 GGGAGACCCAGGCTGTGTGACATCTTTGGCGAGTCCGCGGAGCCATAAGTCTTTAGT 815
661 GGGAGACCCAGGCTGTGTGACATCTTTGGCGAGTCCGCGGAGCCATAAGTCTTTAGT 720
816 CTTATATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 875
721 CTTATATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
876 GCATCATCTTACCTGAGAGGCCATGATTAAGAGAGTGAAGACCTGACAGTGGTT 935
781 GCATCATCTTACCTGAGAGGCCATGATTAAGAGAGTGAAGACCTGACAGTGGTT 840
936 GCATCTTGTGTGTAACAATGCTGACAGTCTGAGAGCCCTGTGAGTGGCTGAGAGACA 995
841 GCATCTTGTGTGTAACAATGCTGACAGTCTGAGAGCCCTGTGAGTGGCTGAGAGACA 900
996 AAAACCTTCAAGAGAGTGTGACCTTCAAGCCAGAGAAAAAAGTCTTCACTGAGTGGTT 1055
901 AAAACCTTCAAGAGAGTGTGACCTTCAAGCCAGAGAAAAAAGTCTTCACTGAGTGGTT 960
1056 GATGGTCTTTCTTCTTATGAGCTCTAGATCTATGCTCAGAGAAAGCATTTAAGCA 1115

961 GATGGTCTTTCTTCTTATGAGCTTATGATCTATGCTCAGAGAAAGCATTTAAGCA 1020
1116 ATTTCTTCATCATGGAGATCAATAACCAAGAGTGGCTTCTGCTGCTATGAAGAG 1175
1021 ATTTCTTCATCATGGAGATCAATAACCAAGAGTGGCTTCTGCTGCTATGAAGAG 1080
1176 GCTCCAGAGATCTGAGTGGCTTCAAGAGTCCCTTCCATCTGATGAAGAAATC 1235
1081 GCTCCAGAGATCTGAGTGGCTTCAAGAGTCCCTTCCATCTGATGAAGAAATC 1140
1236 CTGACATCCCGCTCAGATTTTGAACCTTGTGCTTAATGAATTACTTCATGACAGCAC 1295
1141 CTGACATCCCGCTCAGATTTTGAACCTTGTGCTTAATGAATTACTTCATGACAGCAC 1200
1296 TCCCTGACGAAATCCGAGACAGTCTTCTGAGCTTCTTGGAGATGCTTTGTGTC 1355
1201 TCCCTGACGAAATCCGAGACAGTCTTCTGAGCTTCTTGGAGATGCTTTGTGTC 1260
1356 CCGACATGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTACTTCTATGAG 1415
1261 CCGACATGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTACTTCTATGAG 1320
1416 TTTGGCACCGGCTCAGTGTGTTGAAGACAGAAACCGCTTTTGTCAAAGCCGACAC 1475
1321 TTTGGCACCGGCTCAGTGTGTTGAAGACAGAAACCGCTTTTGTCAAAGCCGACAC 1380
1476 GCTGATGAAGTCCGCTTGTGTTGGTGGTGGCTTCTTGAAGGGGACATTTGATGTC 1535
1381 GCTGATGAAGTCCGCTTGTGTTGGTGGTGGCTTCTTGAAGGGGACATTTGATGTC 1440
1536 GAAAGAGCAGAGAGAGAGAGAGTAACTGAGCCGAGATGATGAATTAAGGGCTACC 1595
1441 GAAAGAGCAGAGAGAGAGAGAGTAACTGAGCCGAGATGATGAATTAAGGGCTACC 1500
1596 TTTGCTCGAACCGGAGATCTATGAGAACGACTGTCTGTGGCAGCTTATTAATCTG 1655
1501 TTTGCTCGAACCGGAGATCTATGAGAACGACTGTCTGTGGCAGCTTATTAATCTG 1560
1656 ACTGAGCAGTACTCTCAGCTGACCTTGAACATGAGCTTGGACAGAGATCAAGAAACCG 1715
1561 ACTGAGCAGTACTCTCAGCTGACCTTGAACATGAGCTTGGACAGAGATCAAGAAACCG 1620
1716 CCGGAGATTTTGAACAGACCAATCCCGGATCTGTGCTCCGACATGCTCCAG 1775
1621 CCGGAGATTTTGAACAGACCAATCCCGGATCTGTGCTCCGACATGCTCCAG 1680
1776 AGTCCCTTTCTTCTTAACTTCTCTCTCTCTCCAGACCTTTCTTTCTTTGTGCT 1835
1681 AGTCCCTTTCTTCTTAACTTCTCTCTCTCTCTCCAGACCTTTCTTTCTTTGTGCT 1740
1836 CCTTGA 1841
1741 CCTTGA 1746

RESULT 3
US-09-799-451-155
Sequence 155, Application US/09799451
Patent No. 6783969

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yunging
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong

Db 637 GCAACTGGGGCTAAGCTGACCAAGTGGCTGCACTACGCTGGTCCAGCAAAATATGCCCC 696
 Qy 748 TCTTGGTGGGGACCCCAAGCTCTGACATCTTTGGCGAGTCCGGGGAGCATTAAGG 807
 Db 697 ACTTTGGAGGCAACCTTGACCGGTGACCAATTTTGGCGAGTCTGGGGTGGACGAGTG 756
 Qy 808 TTTTCACTTATCTACTGCTCTCCCAATGGCCAAAGGCTTATTCACAAAGCCATCATAGAGA 867
 Db 757 TCTCTTCGCTTGTGTGTCTCCCATATCCCAAGAGACTTTCACGAGAGCATATAGAGA 816
 Qy 868 GTGGGGTGGCCATCATCTCTTACCTGGAGGCCCATGATTATGAAAGTGGAGACTGCG 927
 Db 817 GTGGGGTGGCCCTCTGCGCCGCGCTCATTTGGCAGCTCAGCTGA-----TGTATCTCCA 870
 Qy 928 AGTGTGGTGCATATTTCTGTGTGTAACAATGCGTCAAGCTCTGAAGGCCCTGTGAAGTGGC 987
 Db 871 CGGTGTGGCCCAACCTGTCTGCTGTGACCAAGTGTGACTGTAGAGCCCTGTGGGGTGGCC 930
 Qy 988 TGAGGACAAACCCCTCCAGAGAGCTGACCCCTCAGCCAGAAACAAAGTCTTTCACTC 1047
 Db 931 TCCGGGGCAAGATTAAGAGAGATTTCTTGCAATTAACAAGCTTTCAAGATGATCCCG 990
 Qy 1048 GAGTGTGGATGATGCTTTCTTTCTTATGAGCTCTAGATCTATTTGCTCAAGAAACAT 1107
 Db 991 GAGTGTGGATGATGAGGCTTTCTTCTGCGGACACCCCAAGAGCTGTGGCTGTGCGGACT 1050
 Qy 1108 TTAAGCAATTTCTTCTCATATCGAGTCAATTAACAAGAGTGGCTTCTGTGCTCA 1167
 Db 1051 TTCAGCTGTCTCTAGCATTTGTGTGTGTCACAAACATTAATTCGGCTGGCTATCCCA 1110
 Qy 1168 TGA-----AGAGGCTCTGAGATCTCTAGTGGCTCCAAACATGCTCCCTTGGCC 1215
 Db 1111 AGGTCAATGAGATCTATGATATCCCAAGAGAAATGACAGAGAGCTCTCCAGGCTGTCTC 1170
 Qy 1216 TCCATCTGATACAAACATCTGACATCCGCTCAGATTTTGGCACTTTGTGCTAATG 1275
 Db 1171 TCCAGAAATGTTAAGCTGTGATGTGCTCTCTCAATTTGTGTAAGCTGTGAGGAGG 1230
 Qy 1276 AATATCTTCATGACAAAGCACTCCCTGACTGAAATCCGAGACAGTCTTGTGACTGTG 1335
 Db 1231 AGTACATTTGGGAGCAATGGGAGATCCCGACAGCCCTCCAGGCGAGTTCAGAGATGATG 1290
 Qy 1336 GAGATGTGTTCTTTTGTGCTCCCTGACATGATCAAGCTGATATCAAGAGATGCTGTG 1395
 Db 1291 CGGATCTCATGTTTGTGATCTCTGACATCCAAAGTACAAATTTTCAAGT-----TTCCGGG 1347
 Qy 1396 CACTGTCTACTTCTATGAGTTCGGCAACGGGCTCAGTGTCTTGAAGAACAAGGCCGG 1455
 Db 1348 CCCCTGTGTACTTTCTACAGATTCAGCATCAGCCAGCTGGCTCAAGAACATCAGGCCAC 1407
 Qy 1456 CTTTGTCAAAGCCGACAGCTGATGAATCGGCTTTGTGTGGTGTGCTTCTGA 1515
 Db 1408 CCGACATGAAAGGCAAGCATGTGATGAGCTTCTTTGTTCAGAAATTTCTTTGGGG 1467
 Qy 1516 AGGGGGACATTTGATGTTGAAAGAGCAGCGAGAGAGAGAAATTAAGTGAAGCCGAGAGA 1575
 Db 1468 GCAACTCAATTAAT-----CACTGAGAGAGAGAGAGAGCTTAAGAGAGAGAGA 1515
 Qy 1576 TGATGAATATCTGGGCTACTTTGCTGAAACGGGAAATCTTAATGGAGAGAGAGCTGTCTC 1635
 Db 1516 TGATGAAGTACTGGGCAACTTTTGGAGAAATGGAGAACCCCAATGGCGAGGGTGTGCGAC 1575
 Qy 1636 TGTGGCAGCTTATATCTGACTGAGAGTACTTCAAGCTGGAATTGAAGAGCTGTG 1695
 Db 1576 ACTGGCGCTGTTCGACAGAGAGAGCAATACTGAGCTGGAACCTTAACAGCCGTGCGTGG 1635
 Qy 1696 GACAGAGACTCAAGAACCGCGGGTGAATTTTGAAGCAAGCAATCCCCC 1746
 Db 1636 GCGGGCTCTGAAGGCCCAAGGCTCAGTTCTGGAAGAGGGCGCTGCCCC 1686

RESULT 5
 US-09-949-016-555

; Sequence 555, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 555
 ; LENGTH: 2169
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-555

 Query Match 18.8%; Score 406.6; DB 4; Length 2169;
 Best Local Similarity 56.5%; Pred. No. 1,2e-110;
 Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

 Qy 211 GGAACACCAAGCTGGAGTGGATTCAGGGCAGCAAGTCACTGTCTGGGAGCCCTGTGCG 270
 Db 157 GGACCAACACACAGCGGGGAGGTGTGGGAGCTTGTCTATGTGAAGGGCGCCAGCCG 216
 Qy 271 CTGTGAAGTGTTCCTCGGAGTCCCTTGTGCTGCTCCCGCTGGAGATCCCTGCATTTA 330
 Db 217 GGGTCCAACTTCTTGGAAATTCATTTGCCAAGCACCTTAAGTCCGCTGCGCATTTG 276
 Qy 331 CGAACCCGCAAGCTCATGCGCCCTGGAGTAACTTGCAGAAACCACTTCACTTAATT 390
 Db 277 CACCCCTGAGGCCCTGTAATCTTGAAGTGTGTGAGGGATGGAACCACTCCGCGCA 336
 Qy 391 TGTGCTCCGAATCTCAGAGTGGCTGCTTTGA---TCAACATATGCTCAAGTGCATT 447
 Db 337 TGTGTCTACAGAGACTCAACCGAGTGAAGTCAAGATTTCTTAGCCATTCACATGACT 396
 Qy 448 ACCCGAATTTGGAAGTGTGAGAACTGCTCTACCTGAACATTAATGAGCGCTGCCAG 507
 Db 397 TCCCTTCCGATCTCAATGTCTGAGACTGCTTACTTACGATCTACACGCGGCCATTA 456
 Qy 508 CCGATACAGGCTTCAAGCTCCCGTCTTGTGTGTGTGTTCCAGAGGTGCTTCAAGCTG 567
 Db 457 GCATGAAGGCTTAACCTGCGGGTGTGTGTGATTCACGCTGTGCGCTTTGTTTG 516
 Qy 568 GCTCAGCTCATCTTTGATGAGGTCCGCTGCTGCTCTATAGAGAGCTGTGTTGTG 627
 Db 517 GCATGGCTTCTTGTATGATGTGTTCAATGCTGCTGCTTGGAGAACTGTGTGTGTGA 576
 Qy 628 TGTCTCAAGTACCGGCTAGAGAAATTTGTTTCTTACCAATGAGATCAGATGCTCGG 687
 Db 577 TATCTCAAGTACCGGCTAGAGTGTCTGAGCTTTTGAAGCTGAGAGCAAGCAAGCAG 636
 Qy 688 GGAATCGGCTTCAAGAGCAGAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
 Db 637 GCAATCGGCTTCAAGAGCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 696
 Qy 748 TGTGTGGGAGACCCCAAGCTGTGACCATTTTGGGAGTGTGCGGGAGGCAATTAAGT 807
 Db 697 ACTTTGAGGCAACCTTGACCGGTGACCATTTTGGGAGTGTGCGGGAGGCAATTAAGT 756
 Qy 808 TTTTCACTTATCTATCTGTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATAGAGA 867
 Db 757 TGTCTTCTGTTGTGTGTCTCCCATATCCAAAGAGCTTTCAAGAGAGCATATAGAGA 816
 Qy 868 GTGGGGTGGCCATCATCTCTTACTGAGGCCCATGATTAATGAGAAAGTGAAGACTTGC 927

```

Db      817 GTGGCGTGGCCCTCTGCCCCGCGCTCATTTGCAAGCTGAGTGA-----TGTCACTTCCA 870
Qy      928 AGTGTGTTGCAATTTCTGTGTAAATGCTGAGCTCTGAGGCGCTGTGAGTGGC 987
Db      871 CGGTGTGGCCCAACTGTCTGTGTGACCAAGTTGACTCTGAGGCGCTGTGTGGGCTGGC 930
Qy      988 TGAGGACAAACCCCTCCAGAGAGCTGAGCCCTCAGCGACAGAAAACAAAGCTTTCACTC 1047
Db      931 TCGGGGGCAGAGTAAGAGGAGGATCTTGGCAATTAACAACTTTCAAGATGATCCCG 990
Qy      1048 GAGGTGTGATGAGTCTTTCTTTCTTAATGAGCCTTAATGATCTAATGCTTCAAGAACAT 1107
Db      991 GAGTGTGATGAGTCTTTCTTTCTTCCGCGACAGCCCGAGAGCTGTGAGCTGTGCGACT 1050
Qy      1108 TTAAGCAATTCCTTCCATCATCCGAGATCAATTAACAGTGTGGCTTCTGTGCTTA 1167
Db      1051 TTCAAGCTGTCCCTTATGACTTGTGTGTGACCAACAAATGAAATTCGCTGTGCTATCCCA 1110
Qy      1168 TGA-----AGAGGCTCTGAGATCTCACTGAGCTCCAGAGCTCCCAAGTCCCTGGCC 1215
Db      1111 AGGTCAATGAGATCTATGATACCCAGAAAGAAATGACAGAGAGGCTCCAGAGCTGCTC 1170
Qy      1216 TCCATCTGATPACAAATCATCCTGACATCCCGCTCAGTATTTGACCTTTGGCTAATG 1275
Db      1171 TGCAGAAATGTTAACTGCTGTGATGTTGCTCTTACATTTGATGACCTGTGAGGAGG 1230
Qy      1276 AATTCCTTCATGACAGCACTCCCTGACCTGAAATCCGAGACAGTCTTTGGAATTGGA 1335
Db      1231 AGTACATTTGGGGCAATGAGGAGATCCCGAGACCTCCAGAGGCAATTCAGAGATATATG 1290
Qy      1336 GAGATGTGTTCTTTGTGTGCTCCGCTGACATGATCAGAGCTGATTCAGAGATGTGTG 1395
Db      1291 CGGATCTCATGTTTGTGATCTCTGCACTCCAGATACACATTTTCAGT---TTCCCGGG 1347
Qy      1396 CACTGTCTAATCTTCTAATGAGTTTCGACCGGCTCAGTCTTTGGAAGCAGAGAGCGG 1455
Db      1348 CCCCTGTGATCTTCTACAGATTCACAGATCAGCCAGCTGTGCTCAAGAAATCAGGCCAC 1407
Qy      1456 CTTTGTGCAAGCGGACCAAGCTGATGAAATCGCTTTGTGTGTGCTGTGCTCTGTA 1515
Db      1408 CGCACTGAAGGAGACCAATGATGATGATCTTTGTTTCAAGATTTCTTTGGGG 1467
Qy      1516 AGGGGACATTTGATGTTGAAAGAGCAGGAGAGAGAAATTAATGAGCGGAGAA 1575
Db      1468 GCAACTACATTAAT-----CACTGAGAGAGAGAGCACTTAAGCAGAGAA 1515
Qy      1576 TGAATGAATACCTGGCTACCTTTGCTCGAACCGGAATCTTAATGGAACGACTGTCTC 1635
Db      1516 TGAATGAATACCTGGCTACCTTTGCTCGAACCGGAATCTTAATGGAACGACTGTCTC 1575
Qy      1636 TGTGACCAAGCTTATATCTGACTGAGAGTACCTCAGCTGATGAAATGAGCCTG 1695
Db      1576 ACTGCGCGCTGTTCACACAGAGAGAGCAATACCTCAGCTGAACTTACAGCGTGTGG 1635
Qy      1696 GACAGAGACTCAAGAACCGCGGTGTGATTTTGGACCAAGCAATCCCGC 1746
Db      1636 GCGGGCTCTGAAAGCCCAAGAGCTCAGATTCTGGAAGAGCGGTGCCCC 1886

```

RESULT 6

```

US-09-595-682B-27
; Sequence 27, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SU-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258

```

```

; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-595-682B-27

```

```

Query Match      18.8%; Score 406.6; DB 4; Length 2191;
Best Local Similarity 56.5%; Pred. No. 1,2e-110;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

```

```

Qy      211 GGAAACACGAGCTGGATGATTTACAGGCAAGCAATCACTGTCTGGAAAGCCCTGTGC 270
Db      179 GGACCAACACACAGGAGGAGGTGCTGGAGAGTCTTGTCAATGTGAAGGGGCCAATGCC 238
Qy      271 CTGTGAACGTGTCTCTGAGTCCCTTTGCTGCTCCCGCTGGAGATCCCTGGATTTA 330
Db      239 GGGTCCAAACCTTCTGGGAATTCATTTGCAAGCAACCTTAAGTCCGCTGGATTTG 238
Qy      331 CGAAACCGGACGCTGACATCCGCTGGAGATTAATTCGAGAAAGCAACCTCTTACCTTAAT 390
Db      299 CACCCCTGAGCCCTCTGAATCTTGTGAGTGTGAGAGATGAAACCAACCATCCGAGCA 358
Qy      391 TGTGCTTCAGAACTCAGAGTGTGCTCTTGA---TCAACATGCTCAAGTGCAT 447
Db      359 TGTGCTTACAGAACTCAGCGAGTGAATCAGAGTTCTTACCACTTACATATACCT 418
Qy      448 ACCGAAATTCGAGTGTCAAGAAATGCTCTTACCTGAACATCTAATGCGCTGCGCAG 507
Db      419 TCCCTTCGAACTCAATGTGAGAGACTGCTGTACTCAAGCATCTACAGCGCGCCATA 478
Qy      508 CCGATACAGGCTCAAGCTCCCGCTTGTGTGTTCCAGAGAGTGCCTTCAAGACTG 567
Db      479 GCATGAAAGGCTTAACCTGCTGATGTGTGATTCACAGTGTGTGCTGTTTGTG 538
Qy      568 GCTACGCTCATCTTTGATGAGTCCGCTGTGCTGCTTAATGAGACGCTGTGTGTG 627
Db      539 GCAATGCTCTTGTATGATGATGTTCCATGCTGCTGCTTGAAGACGTGTGTGTA 598
Qy      628 TCGTCAAGTACCGGCTAAGAAATTTGTTCTTACCAATGGATACGATGCTCGG 687
Db      599 TCAATCAAGTACCGGCTGCTGTGTGCTTGAAGCACTGAGAAACAGCAACCG 658
Qy      688 GGAATGAGGCTTCAAGAGCAAGTGTGCTGTGCTGTGCTGCAAGAAATCATGAGT 747
Db      659 GCAACTGGGGCTACCTGAGCAAGTGTGCTGCACTACGCTGGGTCAAGAAATTCGCC 718
Qy      748 TCTTGGTGGGAGCCCGACGCTGTGACATCTTTGGGAGTCCCGGAGCCATTAATG 807
Db      719 ACTTTGAGAGCAACCTGACGCTGTACCAATTTTGGAGATCTGCGGTGTGACGAGT 778
Qy      808 TTTCAATCTTAATGATGCTTCCATGAGCAAGCTTAATTCACAAAGCATCAATGAGA 867
Db      779 TGTCTGCTGTGTGTGCTTCCCATATCCAAAGACTTTTCAAGAGCATTAATGAGA 838
Qy      868 GTGGGTGGCCATATCCCTTAATCTGAGGCGCATGATTAATGAGAGTGAAGACCTGC 927
Db      839 GTGGGTGGCCCTCTGCGCGGCTCATTTGCACTGAGTGA-----TGTCACTTCCA 892
Qy      928 AGTGTGTTGCAATTTCTGTGTAAATGCTGAGACTGAGAGCCCTGTGAGTGGC 987
Db      893 CGGTGTGGCCAACTGTCTGTGTGACCAAGTTGACTCTGAGGCGCTGTGAGTGGC 952
Qy      988 TGAGACAAACCCCTCCAGAGCTGTGACCTTCAAGCAAGAAACAAAGCTTTCACTC 1047
Db      953 TCGGGGCAAGATTAAGAGAGATCTTGGCAATTAACAGCTTTCAAGATGATCCCG 1012
Qy      1048 GAGTGTGATGAGTCTTTCTTCTAATGAGCTCTAATGATGATGCTCAAGAAAGAT 1107

```

Db 1013 GAGTGTGATGAGGGCTCTTCTCCAGGCAACCCAGAGCTGCTGCTCTGCGCACT 1072
Qy 1108 TTAAGCAATTCCTCCATCCAGAGTCAATTAACAGAGTGTGCTTCCGCTGCTCA 1167
Db 1073 TTGAGCCCTGCTCCATGATGTGTGTCAACAAGTAATTCGCTGCTCAATCCCA 1132
Qy 1168 TGA-----AGGAGGCTCCAGATCTCTGAGTCTGCAACAAAGTCCCTTGGCC 1215
Db 1133 AGGTATGAGATCTATGATACCCAGAGAAATGAGACAGAGGCTCCAGGCTGCTC 1192
Qy 1216 TCCATCTGATACAAACATCTGCAATCCGCTCAAGTATTGCACTTGTGCTTAATG 1275
Db 1193 TGCAGAAATGTAAAGCTGTGATGTGCTCTCAATTTGTGTGATCTGCTGAGAGG 1252
Qy 1276 AATCTTCATGACAGAGCACTCCCTGATGAAATCCGAGACAGTCTTCTGACTTGTG 1335
Db 1253 AGTACATTTGGGACATAGGGGATCCCAAGACCTTCAAGCGAGTTTCAAGAGATGATG 1312
Qy 1336 GAGATGTCTTCTTGTGCTCCCTGCACTGATCAAGTCAAGTATGACAGAGATGCTGTG 1395
Db 1313 CGGACTCATGTGTGTGTGATCCCTGCACTCAAGTACATTTTCAAGT---TTCGGGG 1369
Qy 1396 CACCTGTCTACTTCTATGATGTTTGGGCAACGGCTCAAGTCTTTGAAACAGAACCGG 1455
Db 1370 CCCCTGTACTTCTACAGAGTTCCAGCATCAGCCAGCTGGCTCAAGAACATCAGGCCAC 1429
Qy 1456 CTTTGTCAAAAGCCGACCAAGCTGATGAAATCCGCTTGTGTGCTGTGCTTCTTGA 1515
Db 1430 CGCATATAGGCAACACATGATGATGAGCTTCTTGTGTGCAAGTTCCTTGTGGGG 1489
Qy 1516 AGGGGACATTTGTTTGTGCAAGAGGACCAAGAGAGAGAAATTAATGAGCCGGAAGA 1575
Db 1490 GCAACTACATTAAT-----CACTGAGAGAGAGAGAGAGAGTAAAGAGAGAA 1537
Qy 1576 TGATGAAATCTGAGGCTACTCTTGTCTGAACCGGGAATCTTAATGGAACGACTGTCTC 1635
Db 1538 TGATGAAATCTGAGGCTACTCTTGTCTGAAGAAATGGAACCCCAATGGCAGGGTCTGCAC 1597
Qy 1636 TGTGGCCAGCTTATATCTGACTGAGCACTCTCACTGAGACTTGAACATGAGACCTCG 1695
Db 1598 ACTGCCCCCTGTTCAGCAGAGAGAGCAATCTTCACTGAGCTGAACCTGAGCTGGGTG 1657
Qy 1696 GACAGAGACTCAAGAACCGCGGGTGTGATTTTGTGACCGACGATCCGCC 1746
Db 1658 GCCGGGCTCTGAAGGCCACAGGCTCAAGTTCTGAAAGAGCGCTGCCCC 1708

RESULT 7
US-09-595-682B-20
Sequence 20, Application US/09595682B
Patent No. 6800483
GENERAL INFORMATION:
APPLICANT: Danks, Mary K.
APPLICANT: Potter, Philip M.
APPLICANT: Houghon, Peter J.
TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
FILE REFERENCE: Tumor Cells
CURRENT APPLICATION NUMBER: US/09/595,682B
PRIOR FILING DATE: 2000-01-16
PRIOR APPLICATION NUMBER: 60/075,258
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: PCT/US99/03171
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 1717
TYPE: DNA
ORGANISM: Oryctolagus cuniculus
US-09-595-682B-20

Query Match 16.8%; Score 362.6; DB 4; Length 1717;

Best Local Similarity 54.8%; Pred. No. 1,66-97;
Matches 880; Conservative 0; Mismatches 649; Indels 78; Gaps 5;
Qy 186 GGGCTTCTGCTGAAGGGCCACAGAGAAACAGAGCTGGATGATTCAGAGCAAGCAA 245
Db 62 GGGACCCGTGTGCAACACTGTGTGATGATCTGCAATGCAAGTCTTGGGAAATTC 121
Qy 246 GTCACTGTGAGGAAGCCCTGTGCTGAAAGTTCCTCGAGTCCCTTGTCTGCT 305
Db 122 GTCACTGTGAGGAATTTGACACAGCCCTGCTCTTCTGTGAGTCCCTTCCCAAG 181
Qy 306 CCCCCGTGGAAATCCCTGCAATTTAGAACCCGAGCTGATGCTCCCTGGATTAATTG 365
Db 182 CCCCCCTTGGATCCCTGAGTTTGACACACAGCTGCAAGATCATTTGAGCCAGTG 241
Qy 366 CGAAGGCACTCTTACCTTAATTTGTGCTCCAGAAC-----TCAGAGTGTGCTC 419
Db 242 AAGAACACACCTCTTACCTTCCATGTGCTCCAGAGGAGATACAGGAGATATGCTC 301
Qy 420 TTAGATCAACATGCTCAAGGTGATTAACCCGAATTCGAGTGTCAAGAGTCTGCTC 479
Db 302 TCGAGCTCTTCAACACAGAAAGAAACATCCTCTTAAGTTTCTGAGAGCTGCTT 361
Qy 480 TAACGAAATCTATGCTGCTGCTCCAGGCAAGATACAGGCTCCAGCTCCCTGCTG 539
Db 362 TACGTAATTTTACACCTGCTGACCTGAACAAGAGAGGAGCTCCGCTGATGCTG 421
Qy 540 TGTCTCCAGAGAGTCTTCAAGACTGCTCACTCTTCAATTTGATGAGTCCGCTG 599
Db 422 TGGATCCATGAGAGGTGTCTGATGATGATGATGATGATGATGATGATGATGATG 481
Qy 600 GCTGCTTATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 482 TCTGCTTATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
Qy 660 TTCAACATGAGATCAGATGCTCCGAGAACTGAGCTTCAAGAACAGTGTGCTGCT 719
Db 542 TTCAAGCAGAGATGAGCAGCAGGAGAACTGAGCTTCAAGAACAGTGTGCTGCTG 601
Qy 720 CTGCTCTGAGTCCAGAGAAATCTGATGCTTCTGCTGAGAACCCAGCTCTGTGACATC 779
Db 602 CTGCTCTGAGTCCAGAGAAATCTGATGCTTCTGCTGAGAACCCAGCTCTGTGACATC 661
Qy 780 TTTGGCAGTCCGAGGAGCATAAGTGTCTTATGCTTATGCTTATGCTTATGCTTATG 839
Db 662 TTTGAGAGTCAAGAGAGTCAAGTGTCTTATGCTTATGCTTATGCTTATGCTTATG 721
Qy 840 GGGTATTCACAAAGCCATGAGAGAGTGGTGGCCATATCTTACCTGAGGCC 899
Db 722 AATCTCTTCAATCAGCAATTTCCAGAGTGGGCTGCTCTTCAAGTCTTCAAG 781
Qy 900 CATGATTAAGAGAGTGAAGACCTGAGTGTGCTGCAATTTCTGTGTGAACAATCG 959
Db 782 AAGAACCAAGTCTTGGCTGAGAAATTCATCAGAGTGGGTGTAAACCAACCA 841
Qy 960 TCAAGCTGAGGCGCTGAGTGTGCTGAGCAAAACCTCCAAAGAGTGTGAC- 1018
Db 842 TCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 901
Qy 1019 -----CTCAGCCAGAAACAAAGTCT 1040
Db 902 TTGAATAATTAATGAGCTTATGATCTAATGAGCAACCCAAAGAGAAACAGCGCTTC 961
Qy 1041 TTCACTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1100
Db 962 CTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1021
Qy 1101 AAGATTTAAAGCAATCTTCAATCATGAGATCAATTAACAGAGTGTGCTGCTG 1160
Db 1022 AAGAAATCAACATGCTGCTTCACTGATGATGATGATGATGATGATGATGATGATG 1081
Qy 1161 CTGCTTAAAGAGAGGCTCTGAGTCTCAAGTGTCAACAAAGTCTCTCTCTCTCTCTCT 1209

Db 1082 ATCCCAATGCAATGCTGGGCTATCCACTCTGTGAAGGCAAACTGAGACCAAGACAGCT 1141
 QY 1210 -TTGCCCTCATCTGTATACAAAACATCTGTGACATCCCGCTCACTATTTTGCACCTTGTG 1268
 Db 1142 ACAGAACTCTTGTGAAAGTCTTACCCCATTTGCAATGTCTCTAAGAGACTGACTCCAGTG 1201
 QY 1269 GCTAATG---ATACTTCATAGCAAGACATCCCTGACTGAAATCCGAGACAGTCTTGTG 1325
 Db 1202 GCCACTGAGAAATATTAGAGAGGACAGATGACCTGTCAAAAAGAAAGACTTGTCTTG 1261
 QY 1326 GACTTGTCTGAGATGTGTCTTGTGTGCTCCCTGACTGATCAGAGCTGATATACAGA 1385
 Db 1262 GACATGCTTGAAGATTGTATTGTGTGTCCCATCTGTGAATGTGGCTCGTCAACACAGA 1321
 QY 1386 GATCTGTGTGACCTGTCTACTTCTATGATGATTTGGGACACCGGCTCAAGTCTTTGAAGAC 1445
 Db 1322 GATCTGTGAGCCCACTATATGTATGATGTGATGTGCTCCAGACCTTCTATCAGAC 1381
 QY 1446 ACAGAACCCGCTTTTGTCAAAAGCCGACACGCTGATGATGCTTGTGTGCGTGTG 1505
 Db 1382 ATGAGACCCAAAGACAGATGTAGGGGACCATGAGATGAGATCTTCTGTCTTAGAGCC 1441
 QY 1506 GCTTCTGGAAGGGGACATTTGTATTGTGAAGAGACCAAGAGAGAAAGTTACTG 1565
 Db 1442 CCGTTTTTAA-----AGAGGGTGCCACAGAAAGAGATCAAACTG 1483
 QY 1566 AGCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAAACCGGAATCCATATGGGAAC 1625
 Db 1484 AGCAAGATGTGTGAATACTGCGGCACTTTGTCTAGAAATGGAAATCCCATGGAGAA 1543
 QY 1626 GACTGTCTGTGGGCACTTATATCTGATGAGACAGTCTTCAAGCTTGAATTGAAC 1685
 Db 1544 GGGCTTCTCAATGGCCAGCATGTGACTACAGAAAGTTACTCTGAGATGGAGCAAC 1603
 QY 1686 ATGAGCCTGAGACAGACTCAAGAACCGGGGGATTTTGGAC 1732
 Db 1604 ACCCAGCAGCCAGAAAGTGAAGAACAGAAAGTGGCTTTCTGGAC 1650

RESULT 8

US-09-264-737-3
 / Sequence 3, Application US/09264737A
 / Patent No. 6107549
 / GENERAL INFORMATION:
 / APPLICANT: Peng, Paul C.C.
 / APPLICANT: Ruff, Thomas G.
 / TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
 / FILE REFERENCE: 38-21(10551) RLB3 Pyridine Tolerance
 / CURRENT APPLICATION NUMBER: US/09/264,737A
 / EARLIER FILING DATE: 1999-03-09
 / EARLIER APPLICATION NUMBER: 60/077,377
 / NUMBER OF SEQ ID NOS: 11
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 3
 / LENGTH: 1701
 / TYPE: DNA
 / ORGANISM: Rabbit
 US-09-264-737-3

Query Match 16.5%, Score 357, DB 3, Length 1701,
 Best Local Similarity 54.8%, Pred. No. 7.5e-96,
 Matches 866, Conservative 0, Mismatches 635, Indels 78, Gaps 5,
 QY 214 ACACGAGCTGGGATGATTTAGGGCAAGCAAGTCACTGTGCTGGAAAGCCGTGTGCTG 273
 Db 83 ACATCTGAAAGGGGAAAGTCTCTGGGAAAGTCTGTCACTTAGAAGATTTGACAGCCCG 142
 QY 274 TGAACGTGTCTCTGAGAGTCCCTTTGTGTGCTCCCGCGTGGAGATCCCTGCAATTTACGA 333
 Db 143 TGGCGTCTCTCTGAGAGTCCCTTGTGCGCAAGCCCTCTTGAATCCCTGAGGTTCGAC 202

QY 334 ACCGAGCCTGCATGCGCCGTGGATTAATTGCGAAGAACCACTTCTTAACCTTAATTTGT 393
 Db 203 CACCAAGCTCTGAAATATATGAGCACGTGAAGAACCACTTCTTAACCTTCCATGT 262
 QY 394 GCTCCGAAC-----TCAGAGTGGTCTCTTATGATCAACATGCTCAAGTGACTT 447
 Db 263 GCTCCAGAGACCAATATCAAGGGCATATGTCTCGAGCTCTTCAACCAAGAAAGGA 322
 QY 448 ACCGAAATTTGGAGTGCAGAGACGTGCTCTTACTTAACATCTATGCGCTGCCACG 507
 Db 323 AATATCCCTTTAAGTTTCTGAAGATGCTCTTACCTGAATATTTTACACCCCTCTGACC 382
 QY 508 CCGATPACAGCTCCAAAGCTCCCGCTTGTGTGTGTCCGAGAGGTGCTTCAAGACTG 567
 Db 383 TGACAAAGAGAGGAGCGGTGCGCGTGAATGTGTGATCATGAGAGTGTCTGATGTG 442
 QY 568 GCTGAGCTTCATCTTATATGAGTCCGCTGCTGCTATGAGAGCTGTGTGTG 627
 Db 443 GTGAGCATCAACTATATATGACCTGTGCTTCTGCCCATGAGAACGTGTGTGTGA 502
 QY 628 TGTCTCACTACCGGCTAGAAATATTTGGTTCTTCAACACATGGAGATCAGATCTCCG 687
 Db 503 CCATTCAGTACCGCTGTGGCATCTGGGATTTCTTACACAGAGAGATGACACAGCCGAG 562
 QY 688 GGAATGAGCCTTTCAAGACCAAGTGTGCTCTGTCTGAGTCCAGAAAGATCAGAT 747
 Db 563 GGAATGAGTCACTTGTGACCAAGTGTGCTGCTGCGGTGGTCCAGAACATATTTGCCA 622
 QY 748 TCTTGTGAGGAGACCCAGGCTGTGACCACTTTTGGCCAGTCCCGGAGCCATAGT 807
 Db 623 ACTTTGAGAGGAGACCCAGGCTGTGACCACTTTTGGAGTCAAGAGAGTCAAAAGT 682
 QY 808 TTTTATCTTATATCTGTCTCCATGCGCAAGCTTATTTCCAAAGCATCATAGAGA 867
 Db 683 TCTCTATCTTATATATCCCCCTGACCAAGATCTTTCATATGAGCAATTTCCGAGA 742
 QY 868 GTGAGGTGAGCATATCCCTTACTGAGAGCCATATATGAAAGATGAGACCTGCG 927
 Db 743 GTGAGGTGAGCTCTTCTTCAAGTCTTTCAGAGAACCAAGTCTTGTGCGAGAAA 802
 QY 928 AGGTGTGTGACATTTCTGTGTAAACATGCGTGAAGACTGTAGGCGCTGTGAGTGC 987
 Db 803 TTGCCATCGAAGCTGTGTGTAACCAACACACTCGCTGTATGTGCTCTGCGCC 862
 QY 988 TGAAGCAAAACCTCCAGAGAGCTGTAC----- 1018
 Db 863 AGAAGCAAGAGAAAGTCAATGAGGTGACATTTGAATGAATTTATGTGCTTAGATC 922
 QY 1019 -----CCTAGCCAGAAACAAAGTCTTTCAGTCAAGTGTGTATGTGCTTCT 1068
 Db 923 TAGTTGGCAGCCCAAGAGAACACCGCTTCTGACCATGTGTATGTAGGGGTGCG 982
 QY 1069 TTCTTAATGAGCTCTAATCTATTGTCTCAGAAAGCATTTAAGCAATTCCTTCATCA 1128
 Db 983 TGCCAAAAGCACTTCAAGATTTCTGCAAGAAATAACAACTGTGCTGCCTTAACATG 1042
 QY 1129 TCGAGTCAATTAACAGAGTGTGCTTCTGCTGCTCATGAAAGAGCTCTTGAATCC 1188
 Db 1043 TGGGATCAACAGCAAGTGTGTGCTGATTTATCCCATGCAATGTGGGCTATCCAC 1102
 QY 1189 TCAAGTCTCCAAAGTCCCTT-----GCCCTCATCTGTATACAAACATCC 1236
 Db 1103 TCTTGAAGGCAATGTGACCAAGAACAGCTTACAGAACTTTGTGAAAGTCTTACCCA 1162
 QY 1237 TGCATATCCCGCTCAGATTTTGAACCTTGTGTATG---AATCTTCAATGACAGC 1293
 Db 1163 TTGTAAATGTCTCTAAGAGCTGCTCAAGTGGCACTGAAATATTTAGAGGGACAG 1222
 QY 1294 ACTCCCTGATGAAATCCAGACAGTCTTCTGACTTGTGTGAGATGTGTCTTGTG 1353
 Db 1223 ATGACCCGTGTCAAAAGAAAGACTTGTCTGAGACATGCTTGCAGATTTGTATTGTGTG 1282
 QY 1354 TCCCTGACATGATCAAGCTCTGATATCAAGAGATGCTGTGACCTGTCTACTGTATG 1413


```
Db 1414 CCGGGAGCCGAGAGCCCGCCGCAATGAGTCTCTCGAATCAAGAGCGTCCCTGAG 1473
Qy 726 TGGGTCAGAGAAACATGAGTCTTTCGATGGAGACCCAGCTGTGACCATCTTTGGC 785
Db 1474 TGGGTCAGAGAAACATGAGTCTTTCGATGGAGACCCAGCTGTGACCATCTTTGGC 1513
Qy 786 GAGTCGCGGAGGACCATTAAGTCTTCTAAGTCTTAATCTGTCTCCATGCGCAAGGCTTA 845
Db 1534 GAGAGCGCGGAGGACCGCTCGGTGGGCAATGACCTGTCTCCGCGCCAGCGGGGCTG 1593
Qy 846 TTCCAGAAAGCCATCATGAGAGAGTGGGGTGGCCA 879
Db 1594 TTCCAGAGGCGCGTCTCAAGAGCGGTGCCCA 1627
```

RESULT 11

```
US-09-810-861B-4
; Sequence 4, Application US/09810861B
; Patent No. 6770799
; GENERAL INFORMATION:
; APPLICANT: Mor, Teafir S.
; APPLICANT: Soreq, Hermona
; APPLICANT: Arntzen, Charles J.
; APPLICANT: Mason, Hugh S.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
; FILE REFERENCE: BTI-45
; CURRENT APPLICATION NUMBER: US/09/810,861B
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,440
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 4
; LENGTH: 14446
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (11862)..(12157)
; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector
; OTHER INFORMATION: pTM036. Identity of sequence residues 11862-12157 unknown.
US-09-810-861B-4
```

```
Query Match 6.5%; Score 139.6; DB 4; Length 14446;
Best Local Similarity 51.8%; Pred. No. 5, 1e-30;
Matches 422; Conservative 0; Mismatches 374; Indels 18; Gaps 4;
```

```
Qy 75 TTAATAAAGTACGATCCCATGCGACAGAGGACTTACTTATCTGCTTCAACAATGCTG 134
Db 127 TTACGACGATAGCCATGCTCTCCCGGAGTGTCTGTGCAACGCTTCTGCTGCTTCC 186
Qy 135 TTTTCTGATTTCTGACGCCCCCTGTTGGAGACAGACAGTGGGAAAACCTGGGCTTCT 194
Db 187 CCACCTCTTCTCTCTCTCTCTGCTCTGAGTCTGAGAGAGTGGGGGCTGAGAGGCGGAG 246
Qy 195 GCTGAAGGGCCACAGAGGACACACAGCTGGATGATTCAGGCGAAGCAAGTCACTGTG 254
Db 247 GATGACAGCTGCTGTGTAACGCTGCTGGGGCGGCTGCGGGCAATTCGCTGAAGACC 306
Qy 255 CTGGAGAGCCCTGTGCTGTGTAACGCTGCTCTGAGAGTCCCTTGTGCTGCCCGGCTG 314
Db 307 CCGGGGGGCGCTGTCTCTCTCTCTCTGAGTCTCTCTTCTGCGAGCCAGCCAGT 360
Qy 315 GGATCCCTGCGATTTTGAACCCGCGAGCTGATCGCTTGGGATTAATTGCGAGAGCC 374
Db 361 GAGACCGCGCTTCTTCTGCGACCGGAGCCCAAGCAAGCTTGTGTAGGGGTGTAAGAGCT 420
Qy 375 ACCTCTACCTTAATTTGTGCTTCAGAGTCAAGTCAAGTGTGCTCTTAATTAACAGATG 434
Db 421 ACAACCTTCCAGAGTGTCTGTACCAATATGTGACACCTTAATACCAAGGTTTGAAGGC 480
```

```
Qy 435 CTCAGTGCATTAACCCGAATTTCGAGTGTCAAGAGACTGCTCTTACCTGAACATCTAT 494
Db 481 ACCGAGATGTGAAACCCCAACCGTAGCTGAGCGAGGACTGCTGTACTCAACGCTGTG 540
Qy 495 GCGCTGCGCCACGCGGATACAGGCTTCAAGCTCCCGCTTGTGTGTGTTCCAGAGGT 554
Db 541 ACAACCAATACCCCGGCTCAATACCCCA---CCCTGTCTCTGTGTGATCTAAGGGGT 597
Qy 555 GCGTTCAAGACTGAGTCAAGCTC-----CATCTTGAATGAGTCCGCTGCTGCTAT 608
Db 598 GCGTTTACAGTGGGGCTCTCTCTTGAACGTATGAGTGGCCGCTTCTTGATACAGGC 657
Qy 609 GAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
Db 658 GAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 717
Qy 666 ACATGGGATCAAGATCTCTCGGGGAACTGGGCTTCAAGACACAGTGTGTGTGTGTGT 725
Db 718 CCGGGAGCCGAGAGGCCCGCGGCAATGTGGGTCTCTGATCAGAGGCTGTGCTGAG 777
Qy 726 TGGTCCAGAGAAACATGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 785
Db 778 TGGGTGACAGAGAAACGTGACGCTTGTGGGGTGTGACCGACATCAATGAGCTGTGTGG 837
Qy 786 GAGTCGCGGAGGACCATTAAGTGTCTTAATCTGTCTTCCATGCGCAAGGCTTA 845
Db 838 GAGAGCGCGGAGGCGCGCTCGGTGGGATGACACCTGTCTGTCTCCGCGCCAGCGGGCTG 897
Qy 846 TTCCAGAAAGCCATCATGAGAGAGTGGGGTGGCCA 879
Db 898 TTCCAGAGGCGCGTCTGTGCAAGAGGGTGGCCCA 931
```

RESULT 12

```
US-09-810-861B-5
; Sequence 5, Application US/09810861B
; Patent No. 6770799
; GENERAL INFORMATION:
; APPLICANT: Mor, Teafir S.
; APPLICANT: Soreq, Hermona
; APPLICANT: Arntzen, Charles J.
; APPLICANT: Mason, Hugh S.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
; FILE REFERENCE: BTI-45
; CURRENT APPLICATION NUMBER: US/09/810,861B
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,440
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 5
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (11862)..(12157)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: human acetylcholinesterase gene optimized for
; OTHER INFORMATION: expression in plants
US-09-810-861B-5
```

```
Query Match 6.5%; Score 139.4; DB 4; Length 1725;
Best Local Similarity 55.1%; Pred. No. 1, 2e-30;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;
```

```
Qy 270 CCTGTGAACGTTCTCTGAGTCCCTTGTGCTGCTCCCGCTGGAGTCTGCGATTT 329
Db 175 CCTGTCTGTCTTCTCTGCGATCCCTTGTGCGAGCACCAATGAGACCCCGTCTTT 234
Qy 330 ACGAACCGGAGCGCTGATGCGCTGGGATTAATCTTGGAGAGGACACTTCACTTAAT 389
Db 225 CTGCCACCGGAGCCCAAGCAAGCGCTGTGTCAAGGGGTGTAGACGCTACACCTTCAAGAGT 294
```


COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06106
FILING DATE: 19920722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPM/EAH
TELEPHONE/COMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
PCT-US92-06106-1

Query Match 6.5%; Score 139.4; DB 5; Length 1845;
Best Local Similarity 55.1%; Pred. No. 1.3e-30;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

QY 270 CCGTGAACGCTGTTCTCGAGAGTCCCTTGTGCTGCTCCCGCGTGGATCCCGCATTT 329
Db 175 CCGTCTCTGCTTCTCGGCGATCCCTTGGCGAGCCACCATGGACCCCGCTTT 234
QY 330 ACGAACCCGAGCGCTGCATGCCCTGGGAATCTTGGAGAACCCACTCTCACTTAT 389
Db 235 CTGCCACCGAGGCCAAGACGCTTGTAGAGGGGTGTAGACGCTAACCTTTCAGAGT 294
QY 390 TTGTGCTCCAGAACTCAGAGTGGCTGCTTTAGATCAACACATGCTCAAGGTGATTA 449
Db 295 GTCTGTACCAATATGTGACACCTTATCCAGGTTTGAAGGACCGAGATGTGAAC 354
QY 450 CCGAATTTGAGAGTGCAGAGAGTCCCTTCACTGACATCTATGCCCTGCGCAGCC 509
Db 355 CCGAACCGTGAAGTGAAGAGAGTCCCTTCACTGACATCTATGCCCTGCGCAGCC 509
QY 510 GATACAGGCTCCAGGCTCCCGCTTGTGAGTGTCCAGAGAGTCCCTTCAAGCTGGC 569
Db 415 CCTACATCCCCA---CCCTGTCTCTGTGTGATCTATGGGGTGGCTTTCAGTGGG 471
QY 570 TCAAGCTC-----CATCTTGTATGGGTCCGCTGCTGCTTATGAGAGGTGTGGT 623
Db 472 GCTCTCTCTTGAAGGTATGATGAGTCCGCTTCTTGTATGAGAGAGTGTGTG 531
QY 624 GTGTGCTGCAATACCGGCTAGAAATATTTGATTTCTTACCC---ACATGGATCA 680
Db 532 GTGTGCTAGAACTACCGGCTAGAAATATTTGATTTCTTACCC---ACATGGATCA 680
QY 681 GCTCCGGGAACTGGGCTTCAAGAGCAGGTGCTGTCTGTCTGTCTGTCTGTCTGTCT 740
Db 592 GCGCCGGGCAATGTGGGTCTCTGTATCAGAGGCTGGCCCTGCGAGTGGGTGAGAGAAC 651
QY 741 ATCAGATTCTTCGGTGGGAGCCCAAGCTGTGTGACATCTTTGGCAGTCCGGGAGCC 800
Db 652 GTGGCAGCTTTCGGGAGTGGAGCCCAATCAAGTACGCTGTTGGGAGAACCGGGAGCC 711
QY 801 ATAAAGTTTCTTAATCTTATCTGTCTCCATGAGCCAAAGCTTATTTCCAAAGCATC 860
Db 712 GCTCGGAGGAGCAGCAGTGTGTCTCCCGCCAGCGGGGCTGTTCACAGGGCCGTG 771

QY 861 ATGAGAGTGGGGTGGCCA 879
Db 772 CTGACAGCGGTGCCCA 790

RESULT 15
US-09-949-016-1192
Sequence 1192, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1192
LENGTH: 2158
TYPE: DNA
ORGANISM: Human
US-09-949-016-1192

Query Match 6.5%; Score 139.4; DB 4; Length 2158;
Best Local Similarity 55.1%; Pred. No. 1.5e-30;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

QY 270 CCGTGAACGCTGTTCTCGAGAGTCCCTTGTGCTGCTCCCGCGTGGATCCCGCATTT 329
Db 269 CCGTCTCTGCTTCTCGGCGATCCCTTGGCGAGCCACCATGGACCCCGCTTT 328
QY 330 ACGAACCCGAGCGCTGCATGCCCTGGGAATCTTGGAGAACCCACTCTCACTTAT 389
Db 329 CTGCCACCGAGGCCAAGACGCTTGTAGAGGGGTGTAGACGCTAACCTTTCAGAGT 388
QY 390 TTGTGCTCCAGAACTCAGAGTGGCTGCTTTAGATCAACACATGCTCAAGGTGATTA 449
Db 389 GTCTGTACCAATATGTGACACCTTATCCAGGTTTGAAGGACCGAGATGTGAAC 448
QY 450 CCGAATTTGAGAGTGCAGAGAGTCCCTTCACTGACATCTATGCCCTGCGCAGCC 509
Db 449 CCGAACCGTGAAGTGAAGAGAGTCCCTTCACTGACATCTATGCCCTGCGCAGCC 508
QY 510 GATACAGGCTCCAGGCTCCCGCTTGTGAGTGTCCAGAGAGTCCCTTCAAGCTGGC 569
Db 509 CCGAATCCCCA---CCCTGTCTCTGTGTGATCTATGGGGTGGCTTTCAGTGGG 565
QY 570 TCAAGCTC-----CATCTTGTATGGGTCCGCTGCTGCTTATGAGAGGTGTGGT 623
Db 566 GCTCTCTCTTGAAGGTATGATGAGTCCGCTTCTTGTATGAGAGAGTGTGTG 625
QY 624 GTGTGCTGCAATACCGGCTAGAAATATTTGATTTCTTACCC---ACATGGATCA 680
Db 626 GTGTGCTAGAACTACCGGCTAGAAATATTTGATTTCTTACCC---ACATGGATCA 680
QY 681 GCTCCGGGAACTGGGCTTCAAGAGCAGGTGCTGTCTGTCTGTCTGTCTGTCTGTCT 740
Db 686 GCGCCGGGCAATGTGGGTCTCTGTATCAGAGGCTGGCCCTGCGAGTGGGTGAGAGAAC 745
QY 741 ATCAGATTCTTCGGTGGGAGCCCAAGCTGTGTGACATCTTTGGCAGTCCGGGAGCC 800
Db 746 GTGGCAGCTTTCGGGAGTGGAGCCCAATCAAGTACGCTGTTGGGAGAACCGGGAGCC 805
QY 801 ATAAAGTTTCTTAATCTTATCTGTCTCCATGAGCCAAAGCTTATTTCCAAAGCATC 860
Db 806 GCTCGGAGGAGCAGCAGTGTGTCTCCCGCCAGCGGGGCTGTTCACAGGGCCGTG 865

Qy 861 ATGAGAGTGGGTGGCCA 879
Db 866 CTGCAAGCGGTGCCCCCA 884

Search completed: June 14, 2005, 23:03:30
Job time : 367.298 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 14, 2005, 23:03:41 ; Search time 156.986 Seconds
(without alignments)
10633.187 Million cell updates/sec

Title: US-10-023-515-1
Perfect score: 3872
Sequence: 1 ccacgcgcgcgaataacagc.....aaaaaaaaaaaaaaaaaaaaa 2158

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=x1h
-O=/cgn2.1/USPT spo01/6664091/runat.14062005.133458.9325/app.query.fasta.1.4238
-DB=A.GeneSeq.16Dec04 -QFMT=fasta -SUPRFX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=-1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=6664091@cgn2.1.299@runat.14062005.133458.9325 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOBVERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.GeneSeq.16Dec04:*
1: GeneSeqP19808:*
2: GeneSeqP19908:*
3: GeneSeqP20008:*
4: GeneSeqP20018:*
5: GeneSeqP20028:*
6: GeneSeqP20038:*
7: GeneSeqP20038s:*
8: GeneSeqP20048s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	79.5	581	5	ABB79537 Human car
2	3079	79.5	581	8	ADQ89094 Human uro
3	2915	75.3	575	5	ABP61004 Novel hum
4	2900	74.9	642	5	AAE25025 Human dru
5	2900	74.9	642	8	ABM84314 Human dia
6	2607.5	67.3	581	5	ABP61005 Novel hum
7	2606	67.3	525	6	ADA54807 Human pro
8	2597.5	67.1	581	6	ABU54639 Human NOV
9	2588.5	66.9	618	5	ADR19663 Human dru
10	2454	63.4	469	7	ADB64065 Human pro

11	1985	51.3	542	8	ADP50145	Adf50145 Cat cauxi
12	1985	51.3	542	8	ADP50147	Adf50147 Cat cauxi
13	1780	46.0	356	5	ABP61006	ABP61006 Novel hum
14	1262.5	32.6	561	5	AAE20911	Aae20911 Rat carbo
15	1230	31.8	559	3	AAE58981	AAE58981 Breaet an
16	1230	31.8	550	8	ADP98219	ADP98219 Human int
17	1230	31.8	559	4	AAE31700	AAE31700 Protein e
18	1230	31.8	559	4	ABU53223	ABU53223 Human met
19	1230	31.8	559	6	ABR82851	ABR82851 CES2 re1a
20	1230	31.8	559	6	ABR33960	ABR33960 Human car
21	1230	31.8	559	7	ADD46521	ADD46521 Human pro
22	1221	31.5	306	5	AAE20909	AAE20909 Human car
23	1213.5	31.3	561	7	ADD46519	ADD46519 Rat Prote
24	1191	30.8	607	4	ABU53222	ABU53222 Human met
25	1186.5	30.6	554	8	AD179890	AD179890 Mouse liv
26	1181	30.5	583	4	ABG10273	ABG10273 Novel hum
27	1179.5	30.5	571	3	AA71107	AA71107 Human Hyd
28	1179.5	30.5	571	3	AAU12442	AAU12442 Human PRO
29	1179.5	30.5	571	4	AAE04101	AAE04101 Human gen
30	1179.5	30.5	571	5	AAU83696	AAU83696 Human PRO
31	1179.5	30.5	571	5	ABR84949	ABR84949 Human PRO
32	1179.5	30.5	571	5	ABG64341	ABG64341 Human alb
33	1179.5	30.5	571	6	ABO17886	ABO17886 Novel hum
34	1179.5	30.5	571	6	ABU69108	ABU69108 Human PRO
35	1179.5	30.5	571	6	ABU80843	ABU80843 Human PRO
36	1179.5	30.5	571	6	ABO33809	ABO33809 Novel hum
37	1179.5	30.5	571	6	ABU81140	ABU81140 Human PRO
38	1179.5	30.5	571	6	ABO19424	ABO19424 Human sec
39	1179.5	30.5	571	6	ABU66840	ABU66840 Human PRO
40	1179.5	30.5	571	6	ABU59921	ABU59921 Novel sec
41	1179.5	30.5	571	6	ABU69085	ABU69085 Human PRO
42	1179.5	30.5	571	6	ABO25111	ABO25111 Human sec
43	1179.5	30.5	571	6	ABU82152	ABU82152 Novel hum
44	1179.5	30.5	571	6	ABU67116	ABU67116 Human sec
45	1179.5	30.5	571	6	ABU81549	ABU81549 Human sec

ALIGNMENTS

RESULT 1	ABP79537	standard; protein; 581 AA.
ID	ABP79537	
AC	ABP79537;	
XX		
XX	23-SEP-2002 (first entry)	
XX		
DE	Human carboxylesterase family member 53010.	
XX		
KW	Carboxylesterase; enzyme; human; analgesic; nootropic; antiinflammatory; diagnosis; therapy.	
XX		
XX	Homo sapiens.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..26
FT	Protein	/label=Signal_peptide
FT		27..581
FT		/label=Mature_protein
FT	Domain	44..545
FT		/note="carboxylesterase domain"
FT	Region	125..135
FT		/note="predicted carboxylesterase type-B signature"
FT	Active-site	219..234
FT		/note="predicted carboxylesterase type-B serine active site"
XX		
XX	WO200250256-A2.	
XX		
XX	27-JUN-2002.	
XX		
XX	18-DEC-2001; 2001WO-US049075.	

QY 1596 TTGGTCGAACCGGAGATCTTAATGGAACGACTGTCTGTGGCCAGCTTATATCTG 1655
 DB 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProIaIyrAsnLeu 520
 QY 1656 ACTGAGCACTACCTCCAGCTGGAATTGAACATGAGCTTCGACAGACTCAAGAACCG 1715
 DB 521 ThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuIysGlnPro 540
 QY 1716 GGGGTGGATTTTGAACGACGACCATCCCGTGCCTGCGCCGACATGCTCCAC 1775
 DB 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
 QY 1776 AGTCCTCTTCTTCCCTTAATCTTCTCTCTCTCCACGCTTTCTTTTGTGTCTG 1835
 DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePheCysAla 580
 QY 1836 CCT 1838
 DB 581 Pro 581
 RESULT 2
 ADQ89094 standard; protein; 581 AA.
 AC ADQ89094;
 DT 21-OCT-2004 (first entry)
 DE Human urological disorder related protein 53010 SEQ:46.
 DE urological disorder; uropathic; cytostatic; urinary incontinence;
 KM benign prostatic hyperplasia; human.
 OS Homo sapiens.
 PN WO2004065576-A2.
 PD 05-AUG-2004.
 PF 14-JAN-2004; 2004WO-US000750.
 PR 15-JAN-2003; 2003US-0440318P.
 PR 04-FEB-2003; 2003US-0444783P.
 PR 27-MAR-2003; 2003US-0457901P.
 PR 08-MAY-2003; 2003US-0468752P.
 PR 19-MAY-2003; 2003US-0471614P.
 PR 16-JUN-2003; 2003US-0478742P.
 PR 18-JUL-2003; 2003US-0488529P.
 PR 30-JUL-2003; 2003US-0491156P.
 PR 02-SEP-2003; 2003US-0498594P.
 PR 26-SEP-2003; 2003US-0506332P.
 PA (MILL-) MILLENNIUM PHARM INC.
 PI Karicheci V, Silos-Santiago I, Eliasof SD;
 DR WPI: 2004-562167/54.
 DR N-PSDB; ADQ89093.
 PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
 PT 211 or for identifying a compound capable of treating a urological
 PT disorder or identifying and treating a subject having a urological
 PT disorder.
 PS Claim 1; SEQ ID NO 46; 542pp; English.
 CC The present invention describes the use of polypeptides related to
 CC urological disorders for identifying a compound capable of treating a
 CC urological disorder, identifying a subject having a urological disorder,
 CC or treating a subject having a urological disorder. Also described: (1) a
 CC method for identifying a compound capable of treating a urological
 CC disorder; (2) a method for identifying a subject having a urological

CC disorder; and (3) a method for treating a subject having a urological
 CC disorder. The compound has uropathic and cytostatic activities. The
 CC polypeptides related to urological disorders are useful for identifying a
 CC compound capable of treating a urological disorder, identifying a subject
 CC having a urological disorder, or treating a subject having a urological
 CC disorder. Disorders include urinary incontinence and benign prostatic
 CC hyperplasia. The present sequence represents a human urological disorder
 CC related protein, which is used in the exemplification of the present
 CC invention.
 CC
 SQ Sequence 581 AA;
 Alignment Scores:
 Pred. No.: 8,566-291 Length: 581
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 79.52% Indels: 0
 DB: Gaps: 0
 US-10-023-515-1 (1-2158) x ADQ89094 (1-581)
 QY 96 ATGCCAAGGAGACTTACTTCACTGCTTCAATGAGTGTCTTTCTGATTTCCAGCC 155
 DB 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
 QY 156 CTGTTGGGACACAGACAGAGGGGAAAACTGGGCGCTTGCCTGAAGGGCCACAGAGAAC 215
 DB 21 LeuLeuGlyHisArgGlnTrpGlyIysThrGlyProSerAlaGlnGlyProGlnArgAsn 40
 QY 216 ACCAGGCTGGAGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGGCTGTG 275
 DB 41 ThrArgLeuGlyTrpIleGlnGlyIysGlnValThrValLeuGlySerProAlaProVal 60
 QY 276 AAGGTGTCTCGAGATCCCTTGTGCTGCTCCCGCTGGAGATCCCTGGATTTACGAAC 335
 DB 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuAspPheThrAsn 80
 QY 336 CCGGAGCGTGAATGCGCCGCGGATTAAGTGGAGAAAGCACTCCATTTGTTGTC 395
 DB 81 ProGlnProAlaSerProIleAspAsnLeuArgIuAlaHisSerTyrProAsnLeuGly 100
 QY 396 CTCCAGAACTCAGAGTGGCTGCTTGAATCAACATGCTCAAGTGCATTAACCCGAAA 455
 DB 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuIysValHisTyrProIys 120
 QY 456 TTGGAGTGTCAAGAACTGCTTACTGAACATGTATGCGCTGCCACGCCGATACA 515
 DB 121 PheGlyValSerGlnAspCysLeuIysLeuAsnIleTyrAlaProAlaHisIleAlaAspThr 140
 QY 516 GGCTCCAAAGTCCCGCTTGTGTGTTGCCAGAGAGTGCCTTGAAGTGGCTCAGCC 575
 DB 141 GlySerIysLeuProValLeuValIlePheProGlyGlyAlaPheIysThrIleSerAla 160
 QY 576 TCCATCTTGAATGGTCCGCGCTGCTGCTATAGAGAGCTGCTGTGCTCTCCAG 635
 DB 161 SerIlePheAspGlySerAlaLeuAlaIaIaTyrGlnAspValLeuValValValGln 180
 QY 636 TACCGGCTAGGAATATTTGTTCTTCAACATGAGATCAGCATGCTCCGGGAACTGG 695
 DB 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTrp 200
 QY 696 GCGTTCAAGAGACAGAGTGGCTGCTGTGCTCGGCTCGAAGAAATATGAGATTTTGGT 755
 DB 201 AlaPheIysAspGlnValAlaIaLeuSerTrpValGlnIysAsnIleGlnPheGly 220
 QY 756 GGGAGCCCAAGCTGTGACCATCTTTGGCAGATCCGCGGAGGCATAGTGTCTTAGT 815
 DB 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
 QY 816 CTTAATAGTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGTG 875
 DB 241 LeuIleLeuSerProMetAlaIysGlyLeuPheHisIysAlaIleMetGlnSerGlyVal 260

[illegible]

ABP61004;
10-SEP-2002 (first entry)
Novel human protein. SEQ ID 91.
Human, cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
neurotic; neuroprotective; immunosuppressive; haemostatic;
antiinflammatory; cardiant; antitumor; virucide; antihypertensive;
cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
wound healing disorders; atherosclerosis; Parkinson's disease;
Alzheimer's disease; autoimmune disorder; hematopoietic disorder;
inflammation; neoplastic disease; nervous system disorder;
cardiovascular disorders; pancreatitis; respiratory disorder;
hyperproliferation; systemic autoimmune disease; hyper-immunity;
developmental abnormality; gastrointestinal ulceration; neuropathy;
hematological disease; metabolic disease; sperm dysfunction;
cone photo-transduction deficiency; neurological disease; stroke;
trachea; thymus; lymph node; muscular system; obesity; anorexia;
growth abnormality; precocious puberty.
Homo sapiens.
WO200250105-A1.
27-JUN-2002.
17-DEC-2001; 2001WO-US049322.
19-DEC-2000; 2000US-0256710P.
20-DEC-2000; 2000US-0257048P.
09-JAN-2001; 2001US-0260482P.
30-JAN-2001; 2001US-0264922P.
06-FEB-2001; 2001US-0266797P.
19-MAR-2001; 2001US-0276988P.
04-APR-2001; 2001US-0281535P.
08-MAY-2001; 2001US-0289622P.
(SMK) SMITHKLINE BEECHAM CORP.
(SMK) SMITHKLINE BEECHAM PLC.
(GLAX) GLAXO GROUP LTD.
Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;
Marensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
N-PSDB; ABO86169.
Secreted proteins and polynucleotides useful as vaccines for preventing
or treating various diseases e.g. cancer, wounds, atherosclerosis,
Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder
Claim 1(a); Page 312-313; 335pp; English.

proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;
 aetima; neurological disorder; Alzheimer's disease; Huntington's disease;
 dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;
 drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;
 renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;
 anoxemia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;
 goltre; gastrointestinal disorder; gene therapy; viricide; anticoagulant;
 anticonvulsant; nootropic; enzyme; DME-10.

Homo sapiens.

Key Location/Qualifiers
 Peptide 1..50
 /label= Signal_peptide
 Domain 32..56
 /note= "Transmembrane domain"
 Protein 51..642
 /note= "Mature human DME-10"
 Domain 113..135
 /note= "Transmembrane domain"
 Domain 204..220
 /note= "Transmembrane domain"
 Domain 234..250
 /note= "Transmembrane domain"
 Domain 287..314
 /note= "Transmembrane domain"
 Domain 463..491
 /note= "Transmembrane domain"

MO200246426-A2.

13-JUN-2002.

04-DEC-2001; 2001MO-US047429.

08-DEC-2000; 2000US-0254308P.
 15-DEC-2000; 2000US-0256189P.
 21-DEC-2000; 2000US-0257133P.
 19-JUN-2001; 2001US-0262706P.
 02-FEB-2001; 2001US-0266020P.

(INCY-) INCYTE GENOMICS INC.

Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,
 Lee EA, Ding L, Hafalla AJA, Tang YT, Yue H, Tribouley CM, Lu DM,
 Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y,
 Ison CH,
 WPI; 2002-519668/55.
 N-PSDB; AAD40574.

Novel human drug metabolizing polypeptide, useful in diagnosis,
 prevention or treatment of autoimmune/inflammatory, cell proliferative,
 neurological, developmental, endocrine, metabolic and gastrointestinal
 disorders.

Claim 65; Page 155-156; 169pp; English.

The invention relates to an isolated human drug metabolizing enzyme (DME)
 and its nucleotide. DME is useful for diagnosing, treating or preventing
 disorders associated with aberrant expression of DME, where the disorders
 are selected from autoimmune/inflammatory disorder such as acquired
 immunodeficiency syndrome (AIDS), asthma, arteriosclerosis, psoriasis,
 uveitis; a cell proliferative disorder such as arteriosclerosis,
 cirrhosis, hepatitis, and cancer; a neurological disorder such as
 Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;
 anemias; an endocrine disorder such as adenoma, thrombosis and infections
 ; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic
 disorder such as cystic fibrosis, diabetes and goltre; a gastrointestinal
 disorder such as anorexia, peptic ulcer; and liver disorders. DME is
 useful in a number of drug screening techniques and to analyse the
 proteome of a tissue or cell type. The invention is useful for creating

CC knock-in humanised animals or transgenic animals to model human diseases,
 CC in somatic or germline gene therapy, to generate a transcrip image of a
 CC tissue or cell type, for detecting differences in the chromosomal
 CC location due to translocation, inversion, etc. among normal, carrier or
 CC affected individuals, and as hybridisation probes for mapping naturally
 CC occurring genomic sequences. The present sequence is human DME-10
 XX
 SQ Sequence 642 AA;

Alignment Scores:
 Pred. No.: 2 74e-273 Length: 642
 Score: 2800.00 Matches: 548
 Percent Similarity: 99.82% Conservative: 2
 Best Local Similarity: 99.46% Mismatches: 1
 Query Match: 74.90% Indels: 0
 DB: 5 Gaps: 0

US-10-023-515-1 (1-2158) x AAE25025 (1-642)

QY 186 GGGCTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTTACAGGGCAACAA 245
 Db 92 GTPProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValGln 111

QY 246 GTCACTGTGGTGGGAAAGCCCTGGTCCCTGTGAACGTCTTCCGAGTCCCTTTGCTGCT 305
 Db 112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 131

QY 306 CCCCCTGGGATCCCTGCAGATTACGAAACCCGAGCTGCATCGCCCTGGGTAACCTTG 365
 Db 132 ProProLeuGlySerLeuAlaGlnPheThrAsnProGlnProAlaSerProThrPheAsnLeu 151

QY 366 CGAAGAACCACTCCCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGAT 425
 Db 152 ArgGlnAlaThrSerTyrProAlaLeuGlyLeuGlnAsnSerGlnTrpLeuLeuAsp 171

QY 426 CAACACATGCTCAAGGTCATTACCCGAATTCGAGTGCAGAGACTGCGCTCACTG 485
 Db 172 GlnHisMetLeuLysValHisTyrProLysPheGlyValSerGlnAspGlyLeuValTrp 191

QY 486 AACATCTATGCGCTGCGCCACGCGCGAATACAGCTCCAGCTCCCGCTTGTGTGTTCC 545
 Db 192 AsnIleTyrAlaProAlaHisAlaAspTrpGlySerLeuProValLeuValTrpPhe 211

QY 546 CCAGAGAGTCTCTCAAGACTGCTCAGCGCTCATCTTTGATGGTCCGCTGCTGCC 605
 Db 212 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 231

QY 606 TATGAGACGCTGCTGTTGTGCTGCTCAGTACCGGCTAGAAATATTGGTTCTTCACC 665
 Db 232 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 251

QY 666 ACATGGAGTCAGATGCTCCGCGGAACTGGGCTTCAGAGACACAGTGGCTGCTGCC 725
 Db 252 ThrTrpAspGlnHisIleAlaProGlyAsnTrpAlaPheLysAspGlnValAlaLeuSer 271

QY 726 TGGGTCCAGAGAAACATCGAGTTCTTCGGTGGGAGCCAGCTGTGTGACCATCTTGGC 785
 Db 272 TrpValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 291

QY 786 GAGTCGCGCGGAGCCATAGTGTCTTCTGATCTTACTGTCTCCAGTGGCCAAAGCTTA 845
 Db 292 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPrometAlaGlyLeu 311

QY 846 TTCACAAAGCCATCATGAGAGAGTGGGAGCATCATCCCTTACCTGAGGCCCATGAT 905
 Db 312 PheHisLysValAlaIleMetGluSerGlyAlaIleIleIleProTyrLeuGlnAlaHisAsp 331

QY 906 TATGAGAAAGTGAAGGACCTGAGGTTGTCACATTTCTGTGTAACAATGCTGTACAC 965
 Db 332 TyrGlnLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAlaAspSerAsp 351

QY 966 TCTGAGGCTCTGTGAGGTCCTGAGGACAAACCTCCAAAGAGCTGCTGACCTCAGC 1025

```

Db      352 SerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer 371
QY      1026 CAGAAAACAAGCTTTTCACTGAGGCTGTGATGCGCTTTTCTTCTTAAGAGCCTCA 1085
Db      372 GlnLysThrLysSerPheThrArgValValAerGlyAlaPhePheProAsnGluProLeu 391
QY      1086 GATCTATTGTCTCAGAAACATTTAAAGCAATTCCTTCATCATCGAGTCAATAACAC 1145
Db      392 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyAlaAsnAsnHis 411
QY      1146 GAGTGTGCTTCTGTGCTGCTTCAAGAGGCTCTGAGATCTCTAGTGGCTTCAACAG 1205
Db      412 GlnCysGlyPheLeuLeuProMetLysGluAlaProGluValLeuSerGlySerAsnLys 431
QY      1206 TCCCTTGGCCCTCATCTGTATAGAAACATCTGCAATCCCGCTCAGATTGTGACCTT 1265
Db      432 SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleuHisLeu 451
QY      1266 GTGGCTAATGAACTACTCTCATGACAAAGCACTCCCTGACTGAATAATCCGAGACTTTCTG 1325
Db      452 ValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeu 471
QY      1326 GACTTGTCTTGAGATGTGTCTTGTGTGTCTGCTGACATGATCAAGCTGATATCAAGA 1385
Db      472 AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArg 491
QY      1386 GATGTGTGGGCACTGTCTACTTCTATGAGTTTCCGACCGGCTTCAGTCTTGAAGAC 1445
Db      492 AspAlaGlyAlaProValTyrPheTyrGluPheAsnGlnAsnArgProGlnCysPheGlnAsp 511
QY      1446 ACGAAGCCGGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTGTGTGGTGGT 1505
Db      512 ThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGly 531
QY      1506 GCTTCTCTGAGAGGGGACATTTGATTTTGAAGAGCCAGAGAGAGAGAGAGATTACTG 1565
Db      532 AlaPheLeuLysGlyAspIleValMetPheGlnGlyAlaThrGluGluGluLysLeuLeu 551
QY      1566 AGCCGGAAGATGATGAAATATCTGGGTACTTGTCTGCAACGGGAATCTTATGGGAAC 1625
Db      552 SerArgLysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlnAsn 571
QY      1626 GACTGTCTCTGTGGCCAGCTTATATCTGACTGAGCAGTACTCTCCAGCTGAGCTTGAAC 1685
Db      572 AspLeuPheLeuTrpProAlaTyrAsnLeuThrGlnGlnIleuGlnLeuAspLeuAsn 591
QY      1686 ATGAGCTCTGGAACAGACTCAAGAAACCGCGGTGGATTTTTGAACCAAGCCATCCC 1745
Db      592 MetSerLeuGlyGlnArgLeuLysGluProArgValGluPheTrpHisSerThrIlePro 611
QY      1746 CTGATCTCTGTCTGCTCCGACATGCTCCAGAGTCTCTTCTTCTTAATCTTCTCTCT 1805
Db      612 LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 631
QY      1806 CTCCTCCAGCCTTCTTCTTCTTGTGCTCCT 1838
Db      632 LeuLeuGlnProPhePhePheCysAlaPro 642

```

```

RESULT 5
ABM84114
ID      ABM84114 standard; protein; 642 AA.
XX      ABM84114;
AC      18-NOV-2004 (first entry)
DT      Human diagnostic and therapeutic pprotein SEQ ID NO:4363.
XX      Human diagnostic and therapeutic pprotein SEQ ID NO:4363.
XX      gene therapy; human diagnostic and therapeutic polynucleotide; dthp.
XX      Homo sapiens.
XX      WO2004023973-A2.
PN

```

```

XX      25-MAR-2004.
PD
XX      12-SEP-2003; 2003WO-US028227.
PF
XX      12-SEP-2002; 2002US-0410259P.
PR
XX      12-SEP-2002; 2002US-0410260P.
PR
XX      (INCYTE CORP.
PA
XX      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
PI      Hachtshorne TA, Suchorolski MT, Altus CM, Plets SJ, Elder LV;
PI      Mooney EM, Deleage AM, Panesar IS, Banyille SC, Reddy TP;
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI      Perla CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LI;
PI      Lagace RB, Spiro PA, Stewart EA, Wingrove J, Velt UA, Kirton BS;
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI      Patuty S, Shi X, Suarez CJ;
XX
XX      MPI; 2004-329368/30.
DR      N-PSDB; ACN42766.
DR
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT      in diagnosing a condition, disease or disorder associated with human
PT      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT      in gene mapping.
PS      Claim 27; Page; 190pp; English.
XX
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
CC      selected from one of the 2722 sequences defined in the specification. A
CC      polynucleotide of the invention may have a use in gene therapy. The human
CC      diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
CC      used to diagnose a particular condition, disease or disorder associated
CC      with human molecules, e.g. cell proliferative disorders,
CC      autoimmune/inflammatory disorders, developmental disorder, endocrine
CC      disorder, neurological disorders, gastrointestinal disorders, or
CC      infections caused by virus, bacteria, fungi or parasite. The dthp
CC      molecules may also be used in genetic mapping, in identifying individuals
CC      from minute biological samples; in detecting single nucleotide
CC      polymorphisms, as molecular weight markers, and for somatic or germline
CC      gene therapy. The present sequence represents a dthp protein of the
CC      invention. Note: The sequence data for this patent is not represented in
CC      the printed specification, but was obtained in electronic format directly
CC      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
XX      Sequence 642 AA;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.: 2,74e-273 Length: 642
XX      Score: 2800.00 Matches: 549
XX      Percent Similarity: 99.82% Conservative: 1
XX      Best Local Similarity: 99.64% Mismatches: 1
XX      Query Match: 74.90% Indels: 0
XX      DB: 8 Gaps: 0
XX
XX      US-10-023-515-1 (1-2158) x ABM84114 (1-642)
QY      186 GGGCTTCTGTGAAGGGCCAGAGAAACACAGGCTGGATGATTCAGGCGAAGCA 245
Db      92 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrIleGlnLysGln 111
QY      246 GTCACTGTGCTGGGAAGCCCTGTGCTGGAACGTTCTCTGGAGTCCCTTGTGTGCT 305
Db      112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 131
QY      306 CCCCGCTGGATTCCTGCGATTTTACAAACCGCAGCTGATCCGCTGGATTAATCTG 365
Db      132 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 151
QY      366 CGAAGACCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTTTAGAT 425
Db      152 ArgGluAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTyrLeuLeuLeuAsp 171

```

QY 426 CAACATGCTCAAGGTGATTACCCGAATTGGAGTGTCAAGAACTGCTTAACCTG 485
 DB 172 GlnHmleuLeuYsValHsYrProLySpheGlyValSerGluApsrYsLeuYrLeu 191
 QY 486 AACATGTATGCGCTCCAGCCGCGATACAGGCTCCAGACTCCCGCTGTGTGGTTC 545
 DB 192 AsnIleYrAlAProIleHsValAAspHrGlySerIyLeuProValLeuValTrpHe 211
 QY 546 CCAGAGGCTGCTTCAAGACTGCTGCTCCATCTTTGATGGTCCGCTGCTGCC 605
 DB 212 ProGlyValAlaPheYrThrGlySerAlaSerIlePheApsrGlySerAlaLeuHsAla 231
 QY 606 TATGAGACGTGCTGTGTGTGCTCCAGTCCGCGATGAAATTTGGTTTCTTACC 665
 DB 232 TyrGluApsrValLeuValValValGlnYrHsGlyLeuGlyIlePheGlyPhe***Ser 251
 QY 666 ACATGGATTCAGCATGCTCCGCGGAACTGGGCTTCAAGAACAGTGGCTGTCTGCC 725
 DB 252 ThrTrpApsrGlnHsValAProGlyAAsnTrpAlaPheYsApsrGlnValAlaIleLeuSer 271
 QY 726 TGGGTCCAGAAACATGAGTCTTGGTGGGAGCCGAGCTGTGTGACCATTTGGC 785
 DB 272 TrpValGlnYsAAsnIleGluPhePheGlyGlyApsrProSerSerValIThrIlePheGly 291
 QY 786 GAGTCCCGGAGCCATTAAGTGTCTTACTTACTGTCTCCATGGCCAAAGGCTTA 845
 DB 292 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIyGlyLeu 311
 QY 846 TTCGACAAAGCCATCAGAGAGTGGGAGCCATCATCCCTTAACCTGGAGGCCCATGAT 905
 DB 312 PheIleYsAlaIleMetGluSerGlyValAlaIleIleProYrLeuGluAlaHsApsr 331
 QY 906 TATGAGAAAGTGAAGACTGAGTGTGACATTTCTGTGTGAACAATGGGTGAGAC 965
 DB 332 TyrGluYsSerGluApsrGlnValValAlaHsPheYsGlyAAsnHsAAsrApsr 351
 QY 966 TCGAGGCGCTGAGTGTGCTGAGGACAAACCTCCAAAGAGCTGTGACCTTACG 1025
 DB 352 SerGluAlaLeuLeuArgYsLeuArgTrpYsProSerIyGlyLeuLeuThrLeuSer 371
 QY 1026 CAGAAAACAAGTCTTTCACGAGTGTGAGTGTGCTTCTTCTTCAATAGACTTA 1085
 DB 372 GlnLyThrYsSerPheThrArgValValAAspGlyAlaPhePheProAsnGluProLeu 391
 QY 1086 GATCTATGCTCAGAAAGATTAAAGCAATTCCTTCATATGAGTCAATTAACAC 1145
 DB 392 AspLeuLeuSerGlnYsAlaPheYsAlaIleProSerIleIleGlyValAAsnHs 411
 QY 1146 GAGTGTGCTTCTGCTGCTTGAAGAGGCTCCGAGATCTCAGTGGCTCCAAACAG 1205
 DB 412 GluCyGlyPheLeuLeuProMetYsGluAlaProGluIleLeuSerGlySerAAsnYs 431
 QY 1206 TCCCTTGGCCCTCATGTGATCAAAACATCTGCAATCCGCTCAGTATTTGGACCTT 1265
 DB 432 SerLeuAlaLeuHsIleLeuIleGlnAAsnIleLeuHsIleProProGlnYrLeuHsIleu 451
 QY 1266 GTGGCTAAATGATATCTTCATGACAAAGCACTCCGATGAAATCCGAGACAGTCTTGTG 1325
 DB 452 ValAlaAAsnGluYrPheHsIAspYsHsSerLeuTrpGluIleArgApsrSerLeuLeu 471
 QY 1326 GACTTGTCTGAGATGTGTTCTTTGTGTGCTGCTGACATGATCAGCTCGATATCAGAGA 1385
 DB 472 AspLeuLeuGlyAAspValPhePheValValProAlaLeuIleThrAlaArgYrHsAAsr 491
 QY 1386 GATGTGTGAGACCGTACTCTCATAGTGTGGGACCGGCTCAGTGTCTTGAAGAC 1445
 DB 492 AspAlaGlyAlaProValYrPheYrGluPheArgHsAAsrProGlnYsPheGluApsr 511
 QY 1446 ACGAAGCGGCTTTGTCAAGCGACACGCTGATGAAGTCCGCTTGTGTGCGGTGT 1505
 DB 512 ThrYsProAlaPheValYsAlaAspHsValAAspGluValArgPheValPheGlyGly 531

QY 1506 GCCTTCCTGAAGGGGACATTGTTATGTTGAAAGACCCAGAGAGAGAAATTACTG 1565
 DB 532 AlaPheLeuYsGlyAAspIleValMetPheGluGlyAlaThrGluGluYsLeuLeu 551
 QY 1566 AGCCGGAAGATGATGAATATGAGTACTTGTGCTGAAACGGGGAATCTTAATGGGAC 1625
 DB 552 SerArgYsMetMetYsYrTrpAlaThrPheAlaArgTrpHsAAsnProAsnGlyAAsn 571
 QY 1626 GACTGTCTCTGTGGGACACTTATATGACTGTAGCAAGTACCTTCAGCTGACTTGAAC 1685
 DB 572 AspLeuSerLeuTrpProAlaYrAAsnLeuThrGluGlnYrLeuGlnLeuApsrLeuHs 591
 QY 1686 ATGAGGCTTGAGACAGACTCAAGAAACCGGCGGTGATTTTGGACACAGACATCCCC 1745
 DB 592 MetSerLeuGlyGlnArgLeuYsGluProArgValAAspPheTrpHsSerThrIlePro 611
 QY 1746 CTGATCTGTGTGCTCTCCGACATGCTCCACAGTCTCTTCTTCAACTTCTCTCT 1805
 DB 612 LeuIleLeuSerAlaSerAAspMetLeuHsSerProLeuSerSerLeuThrPheLeuSer 631
 QY 1806 CTCTCCAGCCCTTCTTTCTTTGTGTGCTCT 1838
 DB 632 LeuLeuGlnProPhePhePhePheCysAlaPro 642
 RESULT 6
 ABP61005
 ID ABP61005 standard; protein; 581 AA.
 AC ABP61005;
 XX
 DT 10-SEP-2002 (first entry)
 XX
 DE Novel human protein. SEQ ID 92.
 KW Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;
 KW nootropic; neuroprotective; immunosuppressive; haemostatic;
 KW antinflammatory; cardiac; antidiabetic; antitumor; antitumor;
 KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
 KW wound healing disorders; atherosclerosis; Parkinson's disease;
 KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 KW cardiovascular disorders; pancreatic; nervous system disorder;
 KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
 KW developmental abnormality; gastrointestinal ulceration; neuropathy;
 KW haematological disease; metabolic disease; sperm dysfunction;
 KW thyroid disorder; hypothyroidism; brain damage; colitis;
 KW cone photo- transduction deficiency; neurological disease; stroke;
 KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
 KW growth abnormality; precocious puberty.
 OS Homo sapiens.
 PN MO200250105-A1.
 PD 27-JUN-2002.
 XX
 PF 17-DEC-2001; 2001WO-US049232.
 XX
 PR 19-DEC-2000; 2000US-0256710P.
 PR 20-DEC-2000; 2000US-0257048P.
 PR 09-JAN-2001; 2001US-0260482P.
 PR 30-JAN-2001; 2001US-0264922P.
 PR 06-FEB-2001; 2001US-0266797P.
 PR 19-MAR-2001; 2001US-0276988P.
 PR 04-APR-2001; 2001US-0281535P.
 PR 08-MAY-2001; 2001US-0289622P.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 PA (GLAXO) GLAXO GROUP LTD.
 XX
 PI Agarwal P, Birkeland M, Cogswell JP, Kabinick KF, Lai Y;


```

OY 1269 GCTATGATACTCTTCATGACAGACCTCCCTGACTGAATCCGAGACAGTCTTGAC 1328
DB 386 AIAAAGLTYRPhenylalanylserylserineleuileuileuileuileuileu 405
OY 1339 TTGCTTGAGATGATGTTCTTTGTCCTGACGATGATGATGATGATGATGATGAT 1388
DB 406 LeuLeuGlyAspValPhepheValValProIleLeuIleThrAlaArgyHisArg--- 424
OY 1389 GCTGTGACCTGTCTACTTCTATGAGTTTCGGCACCGGCTCAGTGTCTTGAACACAG 1448
DB 424 ----- 424
OY 1449 AAGCGGCTTTTGTCAAGCGACACGCTGATGAAGTCCGTTGTTCGGTGGTCC 1508
DB 424 ----- 424
OY 1509 TTCCTGAAGGGGACATTGTATGTTTCGAMAGGACCGAGAGAGAGAGATTAAGTAC 1568
DB 425 -----GluGlyIleThrGluGluGluLeuLeuSer 435
OY 1569 CGAAGATGATGAATTAATCTGAGCTTCTGCTGACCGGGAATCTATGGAACGAC 1628
DB 436 ArgLysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 455
OY 1629 CTGTCTGTGGCCGCTTATATCTGACGAGTACCTGACCTGAGCTTGAACATG 1688
DB 456 LeuSerLeuTrpProIleAlaTrpAsnLeuTrpGluGlyIleGluLeuAspLeuAsnMet 475
OY 1689 AGCTCTGACAGACAGCTCAAGAACCGCGGTGATTTTGGACAGACCATCCCTCTG 1748
DB 476 SerLeuGlyGlnArgLeuLeuLysGluProArgValAspPheTrpThrSerThrIleProLeu 495
OY 1749 ATCTGTCTGCTCTCGACAGCTGCTCAAGTCTCTTCTTCTTAACTTCTCTCTCTC 1808
DB 496 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 515
OY 1809 CTCGAGCCCTTCTTTCTTTGTCCTCT 1838
DB 516 LeuGlnProPhePhePhePheCysAlaPro 525

RESULT 8
ABUS4639
ID ABUS4639 standard; protein; 581 AA.
XX
AC ABUS4639;
XX
DT 03-JUN-2003 (first entry)
XX
DE Human NOVX polypeptide #98.
XX
KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atriocentricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200281498-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 06-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.

```

```

PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283513P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.

(CUBA-) CUBAGEN CORP.
XX
FA Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek RA;
XX Patuxian M, Liu X, Gusev VY, Li L, Vernet CM, Zethusen BD;
XX Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
XX Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W;
XX Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
XX MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
XX Ellerman K;
XX MPI: 2003-046858/04.
XX N-PSDB: ABX72267.
XX
DR New isolated NOVX polypeptide useful for treating atherosclerosis,
XX metabolic disorders, diabetes, obesity, infectious disease, anorexia,
XX neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 303; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOVX, and the
XX polynucleotides encoding them. The polypeptides and polynucleotides are
XX useful for diagnosing disease, and screening for potential therapeutic
XX agents. The sequences are useful for treating metabolic disorders,
XX cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
XX stenosis, atrial septal defect (ASD), atriocentricular canal defect,
XX ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
XX septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
XX atherosclerosis, obesity, infectious diseases, anorexia, neurodegenerative
XX disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
XX haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
XX and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides
XX of the invention
XX
SQ Sequence 581 AA;
XX
Alignment Scores:
Pred. No.: 8,966-244 Length: 581
Score: 2597.50 Matches: 497
Percent Similarity: 95.07% Conservative: 4
Best Local Similarity: 94.31% Mismatches: 17

```

Query Match: 67.08% Indels: 9
 DB: 6 Gaps: 3
 US-10-023-515-1 (1-2158) x ABUS4639 (1-581)

```

QY 186 GGGGCTTGTGTGAAGGGCCACAGAGAAACACAGAGCTGGATGATTAAGGCAAGCA 245
DB 27 G1YProSerAlaGluGlyProGlnAArgSerThrArgLeuGlyTrpIleGlnGlyLysGln 46
QY 246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAAGCTGTCTCGAAGTCCCTTTGCTGCT 305
DB 47 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 66
QY 306 CCCCCTGGGAGATCCCTGCGATTACGAACCCGAGAGCTGCATGCGCTGGATTAATTG 365
DB 67 ProProLeuGlySerLeuAArgPheThrAsnProGlnProAlaSerProTrpAsnLeu 86
QY 366 CGAAGAACCCACTCTTCACTTAATTGTGCTCGAAGACTCAGAAGTGGCTCTTAGAT 425
DB 87 ArgGlnAlaThrSerTyProAsnLeuCysLeuGlnAsnSerGluTrpLeuLeuAsp 106
QY 426 CAACACATGCTCAAGGATTAACCGAAATCGAGATGTCAGAAAGCTGCTTCACTG 485
DB 107 GlnHisMetLeuLysValHisTyProLysPheGlyValSerGlnAspCysLeuTyLeu 126
QY 486 AACATCTATGCGCTGCGCCACAGCCGATACAGGCTCCAGGCTCCCGCTTGTGTGCTTC 545
DB 127 AsnIleTyAlaProAlaHisAlaAspThrGlySerTyLeuProValLeuValTrpPhe 146
QY 546 CCAAGAGGCTGCTTCAAGACTGGCTCACTCACTTGTATGGTGGCTGGCTGCTGCC 605
DB 147 ProGlnGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 166
QY 606 TATGAGACGCTGCTGTGTGTGCTGTCAATACCGGCTAGAAATATTGTGTTCTTCAAC 665
DB 167 TyGlnAspValLeuValValValGlnTyArgLeuGlyIlePheGlyPhePheThr 186
QY 666 AACTGGGATCAGATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGCTGTGCC 725
DB 187 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer 206
QY 726 TGGGTCCAGAGAAACATCGAGTTCTTCCGGTGGGAAACCCAGATCTGTGACATCTTTGGC 785
DB 207 TrpValGlnLysAsnIleGlnPhePheGlyLysAspProSerSerValThrIlePheGly 226
QY 786 GAGTCCGCGGAGCCATTAAGTGTCTTAACTGTCTTCCATGCGCAAGGCTTAA 845
DB 227 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 246
QY 846 TTCACCAAAAGCCATGATGAGAGTGGGGGCGCATCCCTTACCTGAGAGCCCATGAT 905
DB 247 PheHisLysValAlaIleMetGlnSerGlyValAlaIleIleProTyLeuGlnAlaHisAsp 266
QY 906 TATGAAAGAGTGAAGACCTTGAGGTGTGTCACATTTCTGTGTAAACATGCGTCAAG 965
DB 267 TyGlnLysSerGlnAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAsp 286
QY 966 TCTGAGGCGCTGCTGAGGTGCTGAGCAAAACCTTCAAGAGCTGCTGACCTCAAC 1025
DB 287 SerGlnAlaLeuLeuAspCysLeuAspTrpThrLysProSerLysGlnLeuLeuThrLeuSer 306
QY 1026 CAGAAAACAAAGTCTTCACTCGATGAGTGTGATGAGTCTTCTTCCATTAAGAGCTCTA 1085
DB 307 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGlnProLeu 326
QY 1086 GATCATATTGTCTCAAGAAAGCATTTAAAGCAATCTTCCATCATCGAGTCAATTAACAC 1145
DB 327 AsnLeuLeuSerGlnLysValAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHis 346
QY 1146 GAGTGTGCTTCTGCTGCTGCTATGAAGAG-----GCTCTGAGACCTCCAGTGGCTCC 1199
DB 347 GlnCysGlyPheLeuLeuProMetValArgIleLeuAlaValHisThrAlaThrProSer 366
  
```

```

QY 1200 AACAACTCCCTTGCCCTC-----CATGTGATCAAAACATCCTGACATC 1244
DB 367 AsnArgAspAlaAlaLeuAlaSerThrAlaGlyHisPheHisValArg-----HisIle 384
QY 1245 CCGCTCACTATTGTCACCTTGTGCTAATGAATCTTCAATGACACACCTCCCTGACT 1304
DB 385 ProProGlnTyLeuHisIleLeuValAlaAsnGlyTyPheHisAspLysHisSerLeuThr 404
QY 1305 GAAATCCGAGACAGTCTTGTGACTTGTGGAGATGTGTTCTTTGTGTGCTCCCTGCACTG 1364
DB 405 GlnIleAspSerLeuLeuAspLeuLeuGlyAspValPhePheValProAlaLeu 424
QY 1365 ATCAACGCTCGATATACAGAGATGCTGTGACACCTGTCTACTTCAATGAATTCGGCAC 1424
DB 425 IleThrAlaArgTyHisAlaArgAspAlaGlyAlaProValTyPheTyGlnPheArgHis 444
QY 1425 CCGCTCACTGCTTGAAGACAGAGCGGCTTTTGTCAAGCCGACACGCTGATGAA 1484
DB 445 ArgProGlnCysPheGlnAspThrLysProAlaPheValLysValAspHisAlaAspGlu 464
QY 1485 GTCCGCTTTGTGTGCTGCTGCTTCCCTGAAGGGGACATTTGTTATGTTCAAGAGGCC 1544
DB 465 ValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlnGlyAla 484
QY 1545 ACGAGAGAGAGAAATTACTAGCCCGAAGATGAATATGAGGCTTACCTTTGCTCGA 1604
DB 485 ThrGlnGlnGlnLysLeuLeuSerArgLysMetCysTyTrpAlaThrPheAlaArg 504
QY 1605 ACCGGAATCCTTAATGGAAGACACCTGCTGTGGGCGACCTTAATACTGATGAGCAG 1664
DB 505 ThrGlnAsnProAsnGlyAsnAspLeuSerLeuTrpAlaTyAsnLeuThrGlnGln 524
QY 1665 TACTCTCACTGCTGACTGAACATGAGCTTGGACAGAGACTCAAGAACCCGCGTGGAT 1724
DB 525 TyLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnProAlaGArgAsp 544
QY 1725 TTTTGACACGACCAATCCCC 1745
DB 545 ValTrpValThrGlyTyPro 551

RESULT 9
ID ADR19663
ID ADR19663 standard; protein; 618 AA.
XX
AC ADR19663;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human drug metabolizing enzyme (DME)-2 protein sequence.
XX
KW drug metabolizing enzyme; DME; cytosolic; immunosuppressive;
KW anti-inflammatory; endocrine; ophthalmological; gastrointestinal;
KW hepatocellular carcinoma; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; endocrine disorder; eye disorder;
KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
KW human.
XX
OS Homo sapiens.
XX
PN MO200226988-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001MO-US030662.
XX
PR 29-SEP-2000; 2000US-0236947P.
PR 06-OCT-2000; 2000US-0238864P.
PR 20-OCT-2000; 2000US-0242323P.
PR 09-NOV-2000; 2000US-0247581P.
PR 16-NOV-2000; 2000US-0249519P.
PR 22-NOV-2000; 2000US-0252834P.
PR 30-NOV-2000; 2000US-0250567P.
XX
  
```

(INCY-) INCYTE GENOMICS INC.

Azizai Y, Baughn MR, Borowsky ML, Ding L, Duggan BM; Elliott VS, Gandhi AR, Griffin JB, Hafalia AVA, Ison CH, Khan FA; Lal P, Lee SA, Lu DAM, Nguyen DB, Arvizu C, Polichy JL, Rankumar J; Ring HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK; Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H; WPI: 2002-362498/39.

N-PSDB; ADR19681.

Human drug metabolizing enzymes, useful in the diagnosis and treatment of disorders associated with aberrant (DME) activity, e.g., cancer and autoimmune disorders.

Claim 1; SEQ ID NO 2; 142bp; English.

This invention relates to novel drug metabolizing enzymes (DME) and the nucleotide sequences which encode them. The invention may be useful for the development of compounds with a cytostatic, immunosuppressive, antiinflammatory, endocrine, ophthalmological, gastrointestinal or hepatotropic activity acting as an agonist or antagonist of drug metabolizing enzyme activity. The invention may be used in the diagnosis and treatment of disorders associated with decreased or increased expression or activity of drug metabolizing enzymes. Such disorders include cancer, cell proliferative disorders, autoimmune/inflammatory, endocrine, eye, gastrointestinal (including liver disorders) and metabolic disorders. The present sequence is that of a human drug metabolizing enzyme (DME) of the invention. Note: This sequence did not form part of the printed specification but was obtained in electronic format from EPO.

Sequence 618 AA;

Alignment Scores:

Pred. No.:	7.03e-243	Length:	618
Score:	2588.50	Matches:	491
Percent Similarity:	94.62%	Conservative:	1
Best Local Similarity:	94.42%	Mismatches:	5
Query Match:	66.85%	Indels:	23
DB:	5	Gaps:	1

US-10-023-515-1 (1-2158) x ADR19663 (1-618)

186 GGGCGCTTGTCTGTAAGGGGCACAGAGAAACACAGAGCTGGATGATTGAGGCAAGCA 245

92 GYProSerAlaGluGlyProGlnAlaGlnThrArgLeuGlyTrpIleGlnGlyArgGln 111

246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAGACGTGTCTCTGGAGTCCCTTGTGCT 305

112 ValThrValIleuGlySerProValProValAlaValIleuGlyValProPheAlaIle 131

306 CCCCCGCTGGGATCCCTGGATTAGAACCCGACCTGCATCGCCCTGGATTAATTG 355

132 ProProLeuGlySerIleuArgPheThrAspProGlnProIleAspProTrpAspAsn 151

366 CGAGAGCCAGCTCTCAACCTAATTGTGGCTCCAGAACTCAGAGTGGCTCTTAGAT 425

152 ArgGlnAlaThrSerTrpProAsnLeuCybLeuGlnAsnSerGluTrpLeuLeuAsp 171

426 CAACACATGCTCAGAGTGATTAACCCGAAATTGAGAGTGCAGAGACTGCTTACTG 485

172 GlnHISmetLeuIlyValHisTrpProIySPheGlyValSerGluAspCybLeuTrpLeu 191

486 AACATCTAAGCCCTGCCCACAGCCGATACAGGCTCCAAGTCCCCGCTTGGTGTGTT 545

192 AsnIleTrpAlaProIleHisAlaAspTrpGlySerIySPheuProValIleuValTrpPhe 211

546 CCAGAGAGGCTCTTCAAGACTGGCTCAGGCTTCATCTTGAATGGATCCGCGCTGGCTGC 605

212 ProGlyGlyAlaPheIySPheTrpGlySerAlaSerIlePheAspGlySerAlaIleAla 231

606 TATVAGAGAGCTGCTGTGTGTGTCGTCAGATACCGGCTAGAGATATTGTGTTCTTACC 665

D	b	222	TyrgIuArValLeValValValValGlntyrGluLeuGlyIlePheGlyPhePheThr	251
O	y	666	ACATGGAGTCAAGCATGTCTCCGGGGAACTGGGCTTTCAGAGCAAGAGTGGCTGTCTGCC	725
D	b	252	ThrtPrArGlnIlnIleValIleProGlyAsnTrpIlePheIysAspGlnValAlaIleLeuSer	271
O	y	726	TGGGTCCAGAAAGACATGAGTCTTCCGTGGGGAACTCCAGCTGTGGACCATCTTTGGC	785
D	b	272	TrpValGlnIysAsnIleGluPhePheGlyGlyIAspProSerSerValThrIlePheGly	291
O	y	786	GAGTCCGGGGAGGCATAGTGTTTCTAGTCTTPrACTGTCTCCATGGCCAAAGGGCTTA	845
D	b	292	GlnSerIleAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu	311
O	y	846	TTCCACAAAGCCATCATGAGAGAGTGGGGTGGCCATCATCCCTTTCCTGGAGCCCATCAT	905
D	b	312	PheIlnIysValaIleMetGluSerGlyValAlaIleIleProGlyrIleGlnIlnIleAsp	331
O	y	906	TATGAAAGAGTGAAGACCTCGAGGTGGTGGACATTTCTGTGGTAAACATGGCGTCAAC	965
D	b	332	TyrGlnIysSerSerGluAspIleGlnIlnValAlaIlnIlePheCysGlyIAsnAsnAlaSerAsp	351
O	y	966	TCTAGAGGCGCTCGAGGTGGCTGAGGCAAAACCCCTCAAGAGAGCTGTACCCCTGAC	1025
D	b	352	SerGlnAlaIleLeuLeuArgCysLeuAlaGthrIysProSerIysGluLeuIleThrLeuSer	371
O	y	1026	CAGAAACAAAGTCTTTTCACTCGAGGTGGTGGTATGGTCTTTCTTTCCTAAATGAGCCTTA	1085
D	b	372	GlnIysThrIysSerPheThrIArgValAlaAspGlyAlaPhePheProAsnGluProIleu	391
O	y	1086	GATCTATTGTCTCAGAAAGACATTTAAAGCAATTCCTTCCATCATCGAGTCAATPACAC	1145
D	b	392	AspIleuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAlaAsnIlnIls	411
O	y	1146	GAGTGTGGCTTCCGTGGCTCTTATGAAGAGGCTCTGAGATCCTCAGTGGGTCCAAACAG	1205
D	b	412	GlnCysGlyPheIleuLeuProMet-----	419
O	y	1206	TCCCTGGCCCTCCATGTGATACAAACATCTGTGACATCCGCGCTCAGTATTTGACACCTT	1265
D	b	420	-----HisIleProProGlnIyrIleuIlnIleu	428
O	y	1266	GTGGCTAAATGAATATCTTCCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTTCTG	1325
D	b	429	ValAlaAsnGlnIyrPheIlnIleAspIlylnIleSerIleuThrGlnIleArgAspSerIleu	448
O	y	1326	GACTTGGCTTGAAGATGTGTTCTTTGTGTGCTCTGACTGATCAACAGCTCCGATATCACAGA	1385
D	b	449	AspIleuLeuGlyAspValPhePheValAlaProIleuIleThrIlnAlaGlyrIlnIlsArg	468
O	y	1386	GATGTGTGTGACSCGTCTATCTTCTAAAGAGTTTGGGACCGGCGCTCAGTGTCTTTGAAGAC	1445
D	b	469	AspAlaGlyAlaProValIyrPheIyrGluPheAsnIArgIlnArgProGlnIAspPheGluAsp	488
O	y	1446	ACGAAAGCCGCTTTTGTCCAAAGCCGACCCACTGATGAAGTCCGCTTGTGTTCGGTGT	1505
D	b	489	ThrIysProIlnPheValIysAlaAspIlnValaAspGlnValArgPheValPheGlyGly	508
O	y	1506	GCGTTCTGAAAGGGGACATTGTATATGTTGCAAGAGCCACGAGAGAGAGAAATTACTG	1565
D	b	509	AlaPheIleuIysGlyAspIlnIleValIleMetPheGlnIylAlaThrGlnIylGlnIylsLeuIeu	528
O	y	1566	AGCGGAAAGATGAAATACTGGGCTTACCTTGTCCGAACCGGGGAATCTTAATGGGAAC	1625
D	b	529	SerIlnGlyIAsnMetLeuIysIyrTrpAlaThrPheAlaIlnGthrIylAsnProAsnGlyIAsn	548
O	y	1626	GACCTGTCTCTGTGGCCAGCTTATATCTGACTGAGACGTACTCCAGCTGAGCTTGAAC	1685
D	b	549	AspIleuSerIleuTrpProIlnAlaIysIlnIleuThrGlnIylGlnIlnIleuAspIleuAsn	568
O	y	1686	ATGAGCTTCGACAGAGCTCAAGAACCGCGGGTGATTTTGGACAGACCATGCCCC	1745

DB 569 MetSerLeuGlyGlnArgLeuYgIuProArgArgAspValTrrpValThrGlyTyrPro 588
RESULT 10
ADB64065
ID ADB64065 standard; protein; 469 AA.
XX
AC ADB64065;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human protein encoded by clone BRAH20021910.
XX
KM Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KM cell regeneration; membrane protein; signal transduction-related protein;
KM transcription-related protein; osteoporosis; neurological disease;
KM cancer; tumour.
XX
OS Homo sapiens.
XX
PN EP1308459-A2.
XX
PD 07-MAY-2003.
XX
PF 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Iinogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
DR WPI; 2003-450961/43.
DR N-PSDB; ADB62095.
XX
PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page; 222p; English.
XX
XX
XX The invention discloses a polynucleotide comprising a sequence selected
XX from 1970 fully defined nucleotide sequences which encode novel
XX polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX or its partial peptide, an antibody binding to the polypeptide or peptide
XX of the polynucleotide, immunologically assaying the polypeptide or
XX peptide of the polynucleotide by contacting the polypeptide or peptide
XX with the antibody of the encoded protein, and observing the binding
XX between the two, a transformant carrying the polynucleotide in an
XX expressible manner, and an antisense polynucleotide. The oligonucleotide
XX is useful as a primer for synthesizing the polynucleotide, or as a probe
XX for detecting the polynucleotide. The polynucleotides and encoded
XX proteins are useful as pharmaceutical agents and many disease-related
XX genes may be included in them, for developing a diagnostic marker or
XX medicines for regulation of their expression and activity, or as targets
XX of gene therapy. The genes are involved in tissue and/or cell
XX regeneration. Membrane proteins, signal transduction-related proteins,
XX transcription-related proteins, disease-related proteins (e.g. osteoporosis,
XX neurological diseases, cancer, tumours. The cDNA may be used to regulate
XX the activity or expression of the encoded protein to treat diseases. The
XX sequence presented is a protein of the encoded protein. Note: Some of the
XX specification data for this patent is not represented in the printed
XX European Patent Office.
XX
SQ Sequence 469 AA;
XX
Alignment Scores: 8.14e-230 Length: 469
Pred. No.:

Score: 2454.00 Matches: 467
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.57% Mismatches: 0
Query Match: 63.38% Indels: 0
DB: 7 Gaps: 0
US-10-023-515-1 (1-2158) x ADB64065 (1-469)
QY 432 ATGCTCAAGTGCATTAACCCGAAATTCGAGAGTCGAGAAAGCTGCTTACGTAAACATC 491
DB 1 MetLeuIysValHisTyrProIysPheGlyValSerGluAspCysLeuTyrLeuAsnIle 20
QY 492 TTGCGCTGCTCCACGCGGATACAGGCTTCCAGCTCCCGCTTGTGTGCTCCACGA 551
DB 21 TyrAlaProIaHisAlaAspThrGlySerIysLeuProValIleuValTrrpPheProGly 40
QY 552 GGTGCTTCAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
DB 41 GlyAlaPheIysThrGlySerAlaSerIlePheAspGlySerAlaIleuAlaIleuValTrrp 60
QY 612 GACGTGCTGT 671
DB 61 AspValIleuValIleuValIleuValGlnTyrArgLeuGlyIlePheGlyPhePheThrTrrp 80
QY 672 GATCAGCATGTCTCCGGGGAATCGGCGCTTCAAGACCAAGTGTGTGTGTGTGTGTGTGT 721
DB 81 AspGlnHisAlaProGlyAsnTrrpAlaPheIysAspGlnValAlaIleuSerTrrpVal 100
QY 732 CAGAAAGACATGAGTCTTGT 791
DB 101 GlnIysAsnIleGluPhePheGlyIleAspProSerSerValThrIlePheGlyGluSer 120
QY 792 GCGGAGCCATTAAGTCTTCTTCACTTAACTGTCTTCCATGCGCAAGGCTTATTCAC 851
DB 121 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeuPheHis 140
QY 852 AAAGCCATCATGAGAGATGGGGGATGCGCATGCTTACCTGAGAGCCATGATTTATGAG 911
DB 141 LysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuIleuAlaHisAspTyrGln 160
QY 912 AAGATGAGGACCTGACAGT 971
DB 161 LysSerGluAspLeuGlnValIleAlaHisPheCysGlyAsnAsnIleSerAspSerGlu 180
QY 972 GCCCTGTGAGGCTGAGGAGCAAAACCTCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1031
DB 181 AlaLeuLeuArgCysLeuArgThrIleAspSerLeuGluLeuLeuThrLeuSerGlnIys 200
QY 1032 ACAAGCTTTTCACTGAGT 1091
DB 201 ThrIysSerPheThrArgValValAlaAspGlyAlaPhePheProAsnGluProLeuAspLeu 220
QY 1092 TTGTCACAAAGCATTTTAAAGCAATTCCTTCATCATGTGAGTCAATTAACAGAGTGT 1151
DB 221 LeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAsnSerHisGluCys 240
QY 1152 GGCTTCTGCTGCTTATGAAAGAGGCTGAGATCCTCAGTGGCTCAAGAGCCCTT 1211
DB 241 GlyPheLeuLeuProMetIysGluIleAspGluIleLeuSerGlySerAsnIysSerLeu 260
QY 1212 GCCCTCATCTATACAAACATCTGTGACATCCGCGCTCATATTTGACCTTGAGCT 1271
DB 261 AlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisIleuValAla 280
QY 1272 AATGAATACCTTCAGACAGACATCCCTGATCTGAATATCCGAGACAGTCTTGTGACTTG 1331
DB 281 AsnGluTyrPheHisAspIysHisSerLeuThrGluIleArgAspSerLeuLeuAspLeu 300
QY 1332 CTGGAGATGT 1391
DB 301 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisIleArgAspAla 320
QY 1392 GGTGACCTGTCTACTTCTATGAGATTTTGGGACCGGCGCTCAGTCTTGAAGACGAGAG 1451

DB 321 GYAlaProValrYrPheTyrgluPheArgHisArgProGluInCysPheGluSerThrLys 340
QY 1452 CCGGCTTTTGTGAAGCCGACACCGCTGATGAAGTCCCTTTGTTCGGTGTGGCTTC 1511
DB 341 ProlaPheValValValaAspHisAlaAspGluValaArgPheValaPheGlyGlyAlaPhe 360
QY 1512 CTGAAGGGGAGCATTTGTTATGTTGAAAGGACCAAGGAGAGAAAGTTACTGAGCCGG 1571
DB 361 LeuLy8GlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuSerArg 380
QY 1572 AAGATGATGAATACTAGGCGCTACCTTTGCTCGAAGCGGAACTCTAATGGAACGACTTG 1631
DB 381 LysMetMetLeuTyTrrPalaThrPheAlaArgThrGlyAsnProGlnGlyAsnAspLeu 400
QY 1632 TCTCTGTGGCCGCTTATATCTGACTGACGAGTACCTCCAGCTGAGCTTGAACATGAGC 1691
DB 401 SerLeuTrrProlaIleTyArgAsnLeuThrGluGlnTyLeuGluLeuAspLeuAsnMetSer 420
QY 1692 CTCGAGACAGACTCAAGAACCGGGGGTGAATTTTGGACGACGATCCCGCTGATC 1751
DB 421 LeuGlyGlnArgLeuLysGluProArgValGluPheTrrPheSerThrIleProLeuIle 440
QY 1752 CTGTCTGCTCCGACATGCTCCAGACTCTCTTTCTTCTTAACCTTCTCTCTCTC 1811
DB 441 LeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeuLeu 460
QY 1812 CAGCCTTTCTTTTCTTTTGTGCTCT 1838
DB 461 GlnProPhePhePheCysAlaPro 469
RESULT 11
ADP50145
ID ADP50145 standard, protein, 542 AA.
XX
AC ADP50145;
XX
DT 12-FEB-2004 (first entry)
XX
DE Cat cauxin protein SEQ ID NO:2.
XX
KW cat; cauxin; cat kidney disease marker; kidney disease.
XX
OS Felis catus.
XX
PN JP2003250575-A.
XX
PD 09-SEP-2003.
XX
PF 04-MAR-2002; 2002JP-00057908.
XX
PR 04-MAR-2002; 2002JP-00057908.
XX
PA (TOHO-) TOHOKU TECHNOARCH KK.
XX
DR WPI: 2004-002277/01.
XX
N-PSDB: ADP50144.
XX
PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
PS Claim 3; SEQ ID NO 2; 33pp; Japanese.

CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.
XX
SQ Sequence 542 AA;
Alignment Scores:
Pred. No.: 5,58e-184 Length: 542
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 51.27% Indels: 2
DB: 8 Gaps: 2
US-10-023-515-1 (1-2158) x ADP50145 (1-542)
QY 186 GGGCTTGTGTGAAGGGCCAGAGGAACACAGGCTGGATTCAGGGCAAGCA 245
DB 22 GlyProlaAlaAspAlaProValaArgSerThrArgLeuGlyTrrPalaArgGlyLysGln 41
QY 246 GTCACTGTGCTGGAGAGCCCTGTGCTGAAAGTGTTCCTGGAAGTCCCTTGTGCT 305
DB 42 ThrThrValLeuGlySerThrValProValaAsnMetPheLeuGlyIleProTyrrAlaIle 61
QY 306 CCCCCGTGGATCCCTCGGATTTACGAACCCGACAGCTGATCCGCTCGGATTAAGTTG 365
DB 62 ProPheLeuGlyProLeuArgPheLysGlnProLysProlaLeuProLysAsnPro 81
QY 366 CGAGAGCCAGCTCTACCCCTAATTGTGGCTCCAGAACTCAGAGAGGCTGCTTAGAT 425
DB 82 ArgAsnAlaThrSerTyrrProLysLeuCyPheGlnAspLeuGlnTrrPheValSerTyrr 101
QY 426 CAACATGCTCAAGGTCATTACCCGAATTCGAGAGTGTGAGAAAGTCTTACCTG 485
DB 102 GlnHisValLeuLysValaArgTyrrProLysLeuGlnAlaSerGluAspCysLeuTyrrLeu 121
QY 486 AACATCTATGCGCCCTGCCACCCGATACAGGCTCCAGAGCTCCCGCTTGTGCTGTC 545
DB 122 AsnIleTyrrAlaProAlaHisAlaAspAsnGlySerAsnLeuProValaMetValTrrPhe 141
QY 546 CCAGAGGTCCTTCAAGCTGCTCAGCTCATCTTGTGAGGGTCCGCTGCTGCTG 605
DB 142 ProGlyGlyAlaPheLysMetGlySerAlaSerPheAspGlySerAlaLeuAlaIle 161
QY 606 TATGAGAGCTGTGCTGTGCTGCTCCAGTACCGGCTAGAAATTTGGTTCTTCAACC 665
DB 162 TyrrGluAspValLeuIleValaThrThrGlnTyrrArgLeuGlyIlePheGlyPhePheAsp 181
QY 666 ACATGGAGTACAGATGCTCCGGGGAATCGGCTTCAAGAGACAGAGTGGCTCTGTCC 725
DB 182 ThrGlyAspGlnHisAlaArgIlyAsnTrrPalaLeuLeuAspGlnValaAlaLeuThr 201
QY 726 TGGGTCCAGAAAGAAATGAGTCTTCTGGTGGGAAACCCAGTCTGTGACATCTTGGC 785
DB 202 TrrPalaArgAspAsnIleGluPhePheGlyAspProArgSerValThrIlePheGly 221
QY 786 GAGTCCGCGGAGGACATAAGTGTTCATGCTTATAGTGTCTCCATGGCCAAAGGCTTA 845
DB 222 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAsnGlyLeu 241
QY 846 TTCACAAAGCCATCATGAGAGAGTGGGTGGCCATCATCCCTTACCTG--GAGGCCAT 902
DB 242 PheHisLysAlaIleMetGlySerGlyValaIleIleProLeuLeuMetArgProPro 261
QY 903 GATTATGGAAGAGTGAAGACCTGAGAGTGTGGAACATTTCTGTGTAACAATGCGTCA 962
DB 262 GlyAspGluArgLysValaAspLeuGlnValaAlaArgIleCysGlyCyHisIleAsp 281
QY 963 GACTGTGAGGCTCTGATGAGTGTGCTGAGAGCAAAACCTCCAAAGAGCTGTGACCTTC 1022


```

Qy 726 TGGGTCGAGAAACATCGAGTTCTTCGGTGGGAGCCCGAGCTCTGTGACCATCTTTGGC 785
Db 202 TTPVAlaIgaApnaIleIgluPhepHeoIglYlaPheProIgaSerValThrIlePheoIgl 221
Qy 786 GAGTCCGGGAGGAGCCATAGTGTCTTAGTCTTACTGTCTCCAGTGGCCAAAGCTTA 845
Db 222 GluSerIalagIAlaIleSerIalSerSerIleuIleuSerProIAlaIamIglYleu 241
Qy 846 TTCCCAAAAGCCATATGAGAGTGGGGTGGCCATCATCCCTTACCTG---GAGGCCAT 902
Db 242 PheIstIyAlaIleMeIGluSerIglYAlaIaIleuProIleuMeIarPPro 261
Qy 903 GATTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
Db 262 GlYapgluIaIglYlaPhepHeuIglYAlaIaIaIglYleCylYCyhIaIaSer 281
Qy 963 GACTCTGAGGCGCTCTGAGAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022
Db 282 AapSerIalalaleuIleuIglYleuIaIglYlaIglYleuIleuIleuIleuIleu 301
Qy 1023 AGCCAGAAACAAAGTCTTCTACTGAGTGGTGTATGGTCTTCTTCTTATAGAGCT 1082
Db 302 SerIyIstIyIleuThrPheSerIleProValIleAapPhePheProIaPPro 321
Qy 1083 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCATTCCTCCATGAGAGAGAGAGAGAG 1142
Db 322 ValAlaIleuIleuIglYlaPhepHeuIleuIleuIleuIleuIleuIleuIleu 341
Qy 1143 CACGAGTGGCTCTCTGCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1202
Db 342 HieIglYsalIaPheIleu---SerThrIgluPheSerIgluIleuIglYleuIleu 360
Qy 1203 AAGTCCCTGGCGCTCTGATGATACAAACATCTGACATCCCGCTCTGATGATGAC 1262
Db 361 ArgSerIleuAlaIleuIglYleuIaIaIleuIleuIleuIleuIleuIleuIleu 380
Qy 1263 CTGTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
Db 381 LeuAlaIaIaPheIstIyIleuIleuIleuIleuIleuIleuIleuIleuIleu 400
Qy 1323 CTGAGACTGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
Db 401 LeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 420
Qy 1383 AGAGATGCTGTCGACTGCTGATCTTATGATGATGATGATGATGATGATGATGAT 1442
Db 421 ArgAapIaIglYAlaIleuIleuIleuIleuIleuIleuIleuIleuIleu 440
Qy 1443 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1502
Db 441 AapThrIaIaPheIstIyIleuIleuIleuIleuIleuIleuIleuIleuIleu 460
Qy 1503 GGTGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1562
Db 461 GlYlaPheIleuIglYlaPheIleuIleuIleuIleuIleuIleuIleuIleu 480
Qy 1563 CTGAGCCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
Db 481 LeuSerIaIglYlaIleuIleuIleuIleuIleuIleuIleuIleuIleu 500
Qy 1623 AAGCAGCTGCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682
Db 501 GluIglYAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 520
Qy 1683 AACATGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
Db 521 SerIalSerIalIglYIleuIleuIleuIleuIleuIleuIleuIleuIleu 540

```

```

AC ABP61006;
XX 10-SEP-2002 (first entry)
DT
XX
XX
DE Novel human protein. SEQ ID 93.
KW Human; cytosolic; vulnary; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antilucer; virucide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
XX
XX
OS Homo sapiens.
PN
XX
XX
PD WO0205105-A1.
XX
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
XX
PR 19-DEC-2000; 2000US-0256710P.
XX
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX
PA (SMK ) SMITHKLINE BEECHAM CORP.
PA (GLAX ) GLAXO GROUP LTD.
XX
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KP, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX
DR WPI; 2002-508784/54.
XX
XX
N-PSDB; ABQ86171.
XX
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX
PS Claim 1(a); Page 315; 335pp; English.
XX
XX

```

The invention relates to an isolated polypeptide with signal sequences which allow it to be secreted extracellularly or membrane associated. The activity of polypeptides of the invention may be described as, CC cytostatic, CC cytotoxic, CC antiarteriosclerotic, CC antiparkinsonian, CC neuroprotective, CC immunosuppressive, CC haemostatic, CC antiinflammatory, CC cardiant, CC antilucer, CC virucide, CC antithyroid, CC cerebroprotective, CC anorectic, CC metabolic, CC vaccine, CC cancer, CC infection, CC wound healing disorders, CC atherosclerosis, CC Parkinson's disease, CC Alzheimer's disease, CC autoimmune disorder, CC haematopoietic disorder, CC inflammation, CC neoplastic diseases, CC nervous system related disorders and CC cardiovascular disorders, CC pancreatitis, CC respiratory disorder, CC hyperproliferation, CC systemic autoimmune disease, CC hyper-immunity, CC developmental abnormality, CC gastrointestinal ulceration, CC neuropathy, CC haematological diseases, CC metabolic diseases, CC sperm dysfunction, CC thyroid disorders e.g. hypothyroidism, CC brain damage, CC colitis, CC cone photo-transduction deficiency, CC neurological diseases, CC stroke, CC angiogenesis,

CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC tracheas, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention

XX Sequence 356 AA;

Alignment Scores:

Pred. No.:	4,556-164	Length:	356
Score:	1780.00	Matches:	340
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	45.97%	Indels:	0
DB:	5	Gaps:	0

US-10-023-515-1 (1-2158) x ABP61006 (1-356)

QY 819 ATACTGCTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGATGGGCTGCC 878
 DB 17 ILeuSerProMetAlaLysGlyLeuPheHISlyeAlaIleMetGlySerGlyValAla 36
 QY 879 ATCATCCCTTACCGGAGGCCCATGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
 DB 37 ILeuLeuProIleuGluAlaHISAspTYGlyLysSerGlyLeuPheGlnValAla 56
 QY 939 CATTTGTGTGATACATGCGTCAAGTCTGAGAGCCCTGCTGAGAGCCCTGAGAGAGAG 998
 DB 57 HisPheCysGlyValAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThrLys 76
 QY 999 CCTCCAGAGAGCGTGTGAGCCCTCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1058
 DB 77 ProSerLysGlyLeuLeuThrLeuSerGlyLysThrLysSerPheThrArgValAlaAsp 96
 QY 1059 GGTGCTTTCTTCTATGAGGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATT 1118
 DB 97 GlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlyAlaPheLysAlaIle 116
 QY 1119 CCTTCATCATGAG 1178
 DB 117 ProSerIleIleGlyValAlaAsnAlaHISGlyGlyPheLeuLeuProMetLysGlnAla 136
 QY 1179 CCTGAGATCTCTGAGTGGCTCAACAAGTCTTGGCCCTGCATGATACAAACATCTCTG 1238
 DB 137 ProGluIleLeuSerGlySerAlaLysSerLeuAlaLeuHISLeuIleGlnAsnIleLeu 156
 QY 1239 CACATCCCGGCTCAGATTATTCAGCTTGTGAGTAAATGATTTCCATGAGAGAGAGAG 1298
 DB 157 HisIleProProGlnIleLysIleValAlaAsnGluTYrPheHISAspLysHisSer 176
 QY 1299 CTGACTGAATCCGAG 1358
 DB 177 LeuThrGluIleLysAspSerLeuLeuAspLeuLeuGlyAspValAlaPheValAlaPro 196
 QY 1359 GCACTGATCAGAGTCTGATATCACAGAGATCTGTGAGAGAGAGAGAGAGAGAGAGAG 1418
 DB 197 AlaLeuIleThrAlaArgTYrHisArgAspAlaGlyAlaProValTYrPheTYrGluPhe 216
 QY 1419 CGGACCGGCGCTCAGTGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
 DB 217 ArgHisArgProGlnCysPheGlnAspThrLysAspProAlaPheValLysAlaAspHisAla 236
 QY 1479 GATGAAGTCCGCTTGTGTTGCTGAGTGGCTCTTGAAGAGAGAGAGAGAGAGAGAGAG 1538
 DB 237 AspGluValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValAlaPheGln 256
 QY 1539 GAGACCGAG 1598
 DB 257 GlyValaThrGluGluGluLysLeuLeuSerArgLysMetCysTYrTrpAlaThrPhe 276
 QY 1599 GCTTGAACCGGAGATCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1658
 DB 277 AlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuIleProAlaTYrAsnLeuThr 296

QY 1659 GAGCAGTACCTCCAGTGGATCTTGAACATGAGCCTCGAGAGAGAGAGAGAGAGAGAGAG 1718
 DB 297 GUGGINTYrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArg 316
 QY 1719 GTGAGATTTTGGACAG 1778
 DB 317 ValAspPheThrPheSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHisSer 336
 QY 1779 CCTTTTCTTCTTAACTTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1838
 DB 337 ProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAlaPro 356
 RESULT 14
 AAE20911
 ID AAE20911 standard; protein; 561 AA.
 AC AAE20911;
 XX 01-JUL-2002 (first entry)
 DT
 XX
 DE Rat carboxylesterase-like enzyme protein.
 XX
 KW Rat; carboxylesterase-like enzyme; organophosphorus intoxication;
 KW osteoporotic; gene therapy; osteoporosis; antisense therapy; cytotoxic;
 KW detoxifying agent; Paget's disease; bone implant degradation; cancer;
 KW dental implant; enzyme; gene expression.
 XX
 OS Rattus sp.
 XX
 PN WO200206454-A2.
 XX
 PD 24-JAN-2002.
 XX
 PE 10-JUL-2001; 2001WO-EP007919.
 XX
 PR 17-JUL-2000; 2000US-0218564P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Xiao Y;
 XX
 DR WPI; 2002-195808/25.
 XX
 PT Novel human carboxylesterase-like enzyme polypeptide, regulators of which
 PT are useful for preventing and treating organophosphorus intoxication,
 PT cancer and osteoporosis.
 XX
 PS Disclosure, Fig 8; 92pp; English.
 XX
 CC The invention relates to a purified human carboxylesterase-like enzyme
 CC polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
 CC screening for agents which decrease or modulate the activity of
 CC carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
 CC useful for treating a carboxylesterase-like enzyme dysfunction related
 CC disease, such as organophosphorus intoxication, cancer and osteoporosis.
 CC Compounds that increase the ability of human carboxylesterase-like enzyme
 CC to bind to organophosphorus compounds are useful as detoxifying agents.
 CC Carboxylesterase-like enzyme agonists and antagonists are useful for
 CC treating osteoporosis. Paget's disease and degradation of bone implants,
 CC particularly dental implants. Carboxylesterase-like enzyme is useful in
 CC diagnostic assays for detecting diseases and abnormalities or
 CC susceptibility to diseases or abnormalities related to the presence of
 CC mutations in the nucleic acid sequences which encode the enzyme. The
 CC coding sequence of carboxylesterase-like enzyme polynucleotide is useful
 CC in gene therapy and for generating antisense oligonucleotides or
 CC ribozymes which specifically bind to mRNA transcribed from
 CC carboxylesterase-like enzyme DNA. These antisense oligonucleotides are
 CC useful for modulating carboxylesterase-like enzyme gene expression. The
 CC present sequence is rat carboxylesterase-like enzyme protein
 XX
 SQ Sequence 561 AA;

Alignment Scores:

Pred. No.: 2,04e-113 Length: 561
 Score: 1262.50 Matches: 275
 Percent Similarity: 62.41% Conservative: 77
 Best Local Similarity: 48.76% Mismatches: 179
 Query Match: 32.61% Indels: 33
 DB: 5 Gaps: 13

US-10-023-515-1 (1-2158) x AAB20911 (1-561)

```

OY 96 ATGCCAAGGGA---CTTACTCATCTGCTTCAAGATGCTTTTCTGATCTCCAG 152
DB 1 MetProAGAGnGlnLeuHisSerTPrLeuAenLalValLeuPheGlyLeuLeuLeu--- 19
OY 153 CCCCTGTTGGACACAGAGATGGGGAACCTGGGCTTCCTGCTGAAGGCCACAGAG 212
DB 20 --LeuLeuLleHleValGln--GlyGlnAerSerProGluSer---SerProLleAerG 36
OY 213 AACACAGGCTGGAGTGTGATGAGGGAAGCAAGTCACTGTGCTGGGAAGCCCTGGCT 272
DB 37 ThrThrHleThrGlyGlnValGlnGlyValLeuAerHleValLlyAerHleThrValaGly 56
OY 273 GTGAACGTGTTCTCGAGCTCCCTTTGCTGCTCCCGCTGGAGATCCCTGCAATTCAG 332
DB 57 ValHleThrHlePheLeuGlyLleProPheAlaLysProProValGlyProLeuAerPheAla 76
OY 333 AACCCGACGCTGCATCGCCCTGGAGTAACTTGCAGAAAGCCCTCTCACTAATTTG 392
DB 77 ProProGluProProGluProTPrSerGlyValArgAerAlaHleSerGlnProAlaMet 96
OY 393 TGCCTCCAGAACTCAAGATGGCTG-----CTCTAGATCAACACATGCTCAAG 440
DB 97 CysLeuGlnAerLeuAerLleLeuAerGlyValGlyLeuAerMetLysMetLle--- 115
OY 441 GTGATTAACCCGAATTCGAGGTGCAAGAACTGCTCTTCACTGAACATTAATGCGCT 500
DB 116 -----LeuSerSerLleSerMetSerGlnAerCysLeuTyLleuAerValTyAlaPro 133
OY 501 GCCACGCCGATACAGAGCTCCCAAGCTCCCGCTTGTGCTGCTCCACAGAGTGGCTTC 560
DB 134 AlaHleValaArgGlyGlySerAerLeuProValMetValTPrLleHleGlyAlaLeu 153
OY 561 AAGACTGGCTCAAGCTCCATCTTTGATGGGTCCGAGCTGGCTGCTTATGAGGAGCTG 620
DB 154 ValValGlyMetAlerMetLysAerGlySerLeuLeuThrValaAerGlyAerLeuVal 173
OY 621 GTTGTGCTGCTCAGTACCGGCTAGAGATATTTGTTTCTTCAACCATGAGATCAGAT 680
DB 174 ValValThrLleGlnTyLleGlyLeuGlyValLeuGlyPhePheSerThrGlyAerGlnHis 193
OY 681 GCTCCGGGGAATGGGACCTTCAAGACCAAGGCTGCTGTCTGGGTCCAGAAAGAC 740
DB 194 AlaArgGlyAerTPrGlyTyLleuAerGlnValaLalAerAerGlyTyValGlnGlnAer 213
OY 741 ATCGAGTTCTTGGTGGGAGCCCAAGCTGTGACATCTTTGGGAGTCCGCGGGAGCC 800
DB 214 LleAlaHleAerPheGlyGlyAerProAerAerValThrLlePheGlyGlySerAerAlaGly 233
OY 801 ATAGATGTTTATGCTTATATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATC 860
DB 234 ThrSerValSerSerHleValLleSerProMetSerGlnGlyLeuPheHleGlyAlaLle 253
OY 861 ATGAGAGTGGGCTGGCCATCATCCCTTACCTGAGAGCCCATGATTTAGAAAGCTAG 920
DB 254 MetGlySerGlyValaLleLeuAerProAerLleSer-----GlnThrSerGln 270
OY 921 GACCTGACG---GTGGTTGCAATTTCTGTGGAACAATGGCGACAGACTGTAGGCGCTG 977
DB 271 ThrValSerThrThrValaLlyAerLysSerGlyCysGlyLleAerMetAerSerGlyAlaLeu 290
OY 978 CTGAGAGTGCCTGAGACAACCCCTCCAGAGAGCTGCTGACCCCTCAGCCAGAAACAAAG 1037
DB 291 ValArgCysLeuAerGlyAlaLysSerGlyAlaGlyLleLeuValLleAerLysValPheLys 310
  
```

```

OY 1038 TCTTTCATCGAGTGTGATGCTGCTTTCTTTCTTATGAGCTTGAATGCTATTTGCT 1097
DB 311 MetLleProAlaValaValaAerGlyGlyPheLeuProAerHleProLysGlyLeuAerAla 330
OY 1098 CAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACCAAGATGTGCTTC 1157
DB 331 SerGlyAerPheHleProValProSerLleLleGlyValaAerHleAerHleTyTyCysCys 350
OY 1158 CTGCTGCTTATG-----AAGAGGCTCTGAGATCTTCATG 1193
DB 351 ThrLleProMetValMetGlyThrAlaGlnLleLleTyGlyLle-----LeuSer 366
OY 1194 GGCCTCAACAAGTCCCTTCCCTCCATCGATACAAACATCCGTGACATCCGCTCAG 1253
DB 367 ArgGlyAerLeuGlnAlaValLeuAerThrAlaAlaGlnMetCysLeuProProGln 386
OY 1254 TATTTGACCTTGTGGCTAATGATCTTCCATGACAAACACCTCCCTGATGAAATCCGA 1313
DB 387 CysGlyAerPheLeuMetCysGlyTyTyMetGlyAerHleThrAerAerSerGlnThrLeuGln 406
OY 1314 GACAGCTTTGGAATGCTTGGAGATGTTCTTTGCTGCTCCCTGACATGATCAGACT 1373
DB 407 LleGlnTyThrGlyMetMetGlyAerPheLeuPheValLleProAlaLeuGlnValaLla 426
OY 1374 CGATATACAGAGATGCTGTCACCTGCTCTTCTTGAAGTTTGGAGCCGCGCTCAG 1433
DB 427 HlePheGlnArg---SerHleAerProValTyPheTyGlyPheGlnHleValaProSer 445
OY 1434 TGCCTTGAAGACACAGAACGCGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTT 1493
DB 446 TyPheLysAerValaArgProHleValaLysValaAerHleAlaAerGlyValaProPhe 465
OY 1494 GTTGTGCTGCTGCTCTTGAAGGGGACATTTGATTTGGAAGACCCAGAGAGAG 1553
DB 466 ValPhe---GlySerPheSerGlyMetLysLeuAerPhe-----ThrGlyGln 481
OY 1554 GAGAGTTCTGAGCCGAGAAATGATGAATCTGGGCTTCACTTGTGCAACCGGGAAAT 1613
DB 482 GlnAerGlyLeuAerSerAerGlyMetCysTyTyTPrAlaAerHleAlaAerGlnGlyAer 501
OY 1614 CCTAATGGGAAGCACTGTCTGTGAGCCAGCTTATCTGATGATGAGGATCACTCCAG 1673
DB 502 ProAerSerGlyGlyLeuProTyTPrProAlaLeuAerHleAerGlyLleTyLeuGln 521
OY 1674 CTGAGCTTGAACATGAGCTCGACAGACATCAAGAAACCGCGGCTGATTTTGGACC 1733
DB 522 LeuAerThrHleProAlaValaAerAerGlyAlaLeuTyValaArgGlyLeuGlnPheTPrHle 541
OY 1734 AGCACCATCCCC 1745
DB 542 TyThrLeuPro 545

RESULT 15
AAB58981
ID AAB58981 standard, protein, 549 AA.
XX
AC AAB58981;
XX
DT 27-MAR-2001 (first entry)
XX
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 689.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW oncogenic; neuroprotective; antiviral; anti-allergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune hemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX
OS Homo sapiens.
  
```

```

XX XX WO2005173-A1.
XX PN
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005881.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-611515/58.
XX DR N-PSDB; AAF21884.
XX PT New human breast and ovarian cancer associated gene sequences and the
XX PT polypeptides encoded by these genes, useful in the prevention, treatment
XX PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
XX PT neurological diseases.
XX PS Claim 11; Page 1142-1144; 1299pp; English.
XX CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX CC associated with breast and ovarian cancer. Included in the invention are
XX CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX CC isolation and characterisation of the DNA and protein sequences of the
XX CC invention. The breast and ovarian cancer associated DNA, protein, agonist
XX CC or antagonist sequences exhibit cytostatic; immunosuppressive; nocotopic;
XX CC neuroprotective; antiviral; antiallergic; hepatocytic; antidiabetic;
XX CC antiinflammatory; anticancer; anticonvulsant; antibacterial;
XX CC protein sequences are used in the diagnosis of cancer, polynucleotide and
XX CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
XX CC and agonists may also be used in the diagnosis, prevention and treatment
XX CC of immune disorders e.g. Addison's disease, allergies, autoimmune
XX CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX CC cardiovascular disorders such as myocardial ischaemia; wound healing;
XX CC neurological diseases such as cerebral anoxia and epilepsy; and
XX CC infectious diseases
XX SQ Sequence 549 AA;

Alignment Scores:
Pred. No.: 3e-110 Length: 549
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 31.77% Indels: 24
DB: 3 Gaps: 8

US-10-023-515-1 (1-2158) x AAB58981 (1-549)
QY 186 GGGGCTTGTGTCAGAGGCGCACAGAAACACAGGCTGGATGGATTACAGGCAAGCA 245
DB 16 GtlylnApsSerAlaSerProIleGrhThrhsthrGlnValLeuGlySerIeu 35
QY 246 GTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCTCGGAGTCCCTTGTGCT 305
DB 36 ValHsValIlysvGlyAlaAbnAlaGlyValGlnThrPheLeuGlyIleProPheAlIly 55
QY 306 CCCCCGTGGAGATCCCTGATATTAACGAACCCGAGCCTGATGCGCCCTGGATTAATTG 365
DB 56 ProIleuGlyProIleuAlaGrPheAlaProProGlnProGlnuSerIyrSerGlyVal 75
QY 366 CGAAGACCACTCTTAATTTGTGCTCCAGAAC-----TCAGAG 410
DB 76 ArgAserGlyThrThrIleProAlaMetCylLeuGlnIlePheLeuThrAlaValGluSerGlu 95
QY 411 TGGGTGCTTAAATCAACACATGCTCAAGTGCAATTAACCGAAATTCGAGTGTCAAA 470
DB 411 TGGGTGCTTAAATCAACACATGCTCAAGTGCAATTAACCGAAATTCGAGTGTCAAA 470

```

```

DB 96 PheIeu-----SerGlnPheAseMetThrPheProSerAspSerMetSerGlu 111
QY 471 GACTGCTTACCTGAACATCTATGCGCTCTGCCACAGCCGATACAGCTCCAAAGCTCCCC 530
DB 112 AspCylLeuIlyrLeuSerIleIyrThrProAlaHsSerHsIleGlyIleSerAbnLeuPro 131
QY 531 GTCTTGCTGTGGTTCCAGAGAGTGCCTTCAAGACTGCTCAGCTCCATCTTGAATGG 590
DB 132 ValMetValITrpIleHsIleGlyIleValLeuValPheGlyMetAlaSerLeuIlyrAbpIly 151
QY 591 TCCGCCCTGGCTGCTATGAGAGCCTGCTGTTGTGTGCTGCATGACCCGCTAGAAATA 650
DB 152 SerMetLeuAlaAlaLeuGlnIleuAbnValValValIleIleGlnIlyrArgLeuGlyVal 171
QY 651 TTTGGTTTCTTCACACACATGAGATGATCAGATGCTCCGGGAACTGGGCTTCAAGACAG 710
DB 172 LeuGlyPhePheSerThrGlyAspIlyHsAlaThrGlyAbnITrpGlyIlyLeuAbpGln 191
QY 711 GTGGCTGCTCTGTCTGTGGGTCCAGAAACATGAGATTTCTCGGTGGAGACCCAGCTCT 770
DB 192 ValAlaAlaLeuAlaGlyITrpValGlnGlnAbnIleAlaHsPheGlyIlyAbnProAspArg 211
QY 771 GTACCATCTTTGGGAGATCCGCGGAGCCATAGTGTCTTATCTTATATCTCTCC 830
DB 212 ValThrIlePheGlyIleuSerAlaGlyIlyThrSerValSerIleuValValSerPro 231
QY 831 ATGCCAAAGGCTTATTTCCAAAGCCATGAGAGATGGGCTGAGCATCATCTTAC 890
DB 232 IleserGlnGlyLeuPheHsIleGlyAlaIleMetGluSerGlyValAlaLeuLeuProGly 251
QY 891 CTGGAGGCCCATGATTTAGAAAGATGAGAGACTG-----CAGTGGTTGCACATTTTC 944
DB 252 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAbnLeu 267
QY 945 TGTGTAAACATGCTGTCAGACTGTGAGGCTCTGAGGCTCTGAGGCTGAGAGCAAAACCTCC 1004
DB 268 SerAlaCylAspGlnValAlaPheSerGlnAlaLeuValGlyCylLeuArgGlyIlySerIly 287
QY 1005 AAGAGCTGCTGACCCCTCAGCCAGAAACAAAGCTTTTCACTGAGTGGTTGATGCTGT 1064
DB 288 GlnIleIleLeuAlaIleAbnIlyProPheIlyMetIleProGlyValAlaPheGlyVal 307
QY 1065 TTTCTTCTAATGAGCTCTAGATCTTATGTCTCAGAAAGCATTTTAAACATTTCTTTC 1124
DB 308 PheLeuProArgHsIleProGlnIleuLeuAlaSerAlaAspPheGlnProValProSer 327
QY 1125 ATCATGCAAGTCAATACCAACGAGTGTGCTTCTGCTGCT-----ATGAAG 1172
DB 328 IleValGlyValAbnAbnAbnGlnPheGlyITrpLeuIleProIlyValMetArgIleIyr 347
QY 1173 GAGGCTCTCGAGATCCTCAGTGGCTCCAAACAGTCCCTTGCCTCATCTGATACAAAC 1232
DB 348 AspThrGlnIlyleuMetAspArgIlyAlaSerGlnAlaAlaLeuGlnIlyMetLeuThr 367
QY 1233 ATCTTGACATCCCGCTGATATTTGACCTTGTGGCTTAATGAATTAATCTCATGACAAG 1292
DB 368 LeuIleuMetLeuProProThrPheGlyAspLeuAlaGlnGlyITryIleGlyAspAbn 387
QY 1293 CACTCCCTGACGAATCCGAGACAGTCTTTCGACTTGTGAGATGTGTTCTTTGTG 1352
DB 388 GlyAspProGlnIThrLeuGlnAlaGlnPheGlnIleuMetAlaAspSerMetPheVal 407
QY 1353 GTCCCTGACATGATCAGACTGATATACAGAGATGCTGGGACCCGCTGACTTCTAT 1412
DB 408 IleProAlaLeuGlnValAlaHsIlePhe---GlnCylSerAlaGlyAlaProValIlyrPheIyr 426
QY 1413 GAGTTTGGCACCGGCTCAAGTCTTGAAGACACGAAGCCGCTTTGTCAAGCCGAC 1472
DB 427 GlnPheGlnIleHsGlnPheSerITrpLeuIlyAbnIleArgProProHsIleMetIlyAlaAsp 446
QY 1473 CACGCTGAATGAATCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1532
DB 447 HsGlyAspGlnIleuProPheValPhe---ArgSerPhePheGlyIlyAbnITryIleIly 465

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen, Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 06:53:49 ; Search time 38.97 Seconds
(without alignments)
10656.177 Million cell updates/sec

Title: US-10-023-515-1
Perfect score: 3872
Sequence: 1 ccacgcgtccgaacacagc.....aaaaaaaaaaaaaaaaaa 2158

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2.1/USPTO_epool/6664091/runat_14062005_133459_9346/app_query.fasta.1.4238
-DB=FIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091 @CGN.1.1_101 @runat_14062005_133459_9346 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298.5	33.5	557	2	A47162
2	1231.5	31.8	532	2	A43329
3	1230	31.8	559	1	JC5408
4	1199	31.0	561	2	S47655
5	1186.5	30.6	554	1	S34607
6	1163	30.0	554	2	A39060
7	1162.5	30.0	566	2	S19307
8	1144	29.5	549	2	JX0054
9	1140	29.4	567	1	A41010
10	1139	29.4	561	2	S62788
11	1138.5	29.4	562	2	A55281
12	1136	29.3	540	2	A31584
13	1129.5	29.2	561	2	S71597
14	1123	29.0	561	2	JC2447

15	1121.5	29.0	565	2	S10367	carboxylesterase (
16	1034.5	26.7	539	2	A29923	carboxylesterase (
17	777	20.1	596	1	ACRYE	acetylcholinestera
18	764.5	19.7	614	2	A39256	acetylcholinestera
19	763	19.7	614	2	JH0314	acetylcholinestera
20	756	19.5	599	1	A38868	acetylcholinestera
21	754	19.5	614	2	JH0811	acetylcholinestera
22	752	19.4	602	1	ACHU	cholinesterase (EC
23	740	19.1	603	2	S70849	cholinesterase (EC
24	732	18.9	581	2	C39768	cholinesterase (EC
25	726.5	18.8	584	2	S48724	acetylcholinestera
26	710.5	18.3	583	2	S10712	acetylcholinestera
27	692.5	17.9	620	2	A54413	acetylcholinestera
28	669.5	17.3	745	2	S13586	triacylglycerol 11
29	658.5	17.0	489	2	B65680	para-nitrobenzyl e
30	643.5	16.6	597	2	A33668	sterol esterase (E
31	633.5	16.4	599	2	A57701	sterol esterase (E
32	627.5	16.2	664	2	JC7990	acetylcholinestera
33	618	16.0	612	2	A34967	sterol esterase (E
34	594.5	15.4	691	2	JE0150	acetylcholinestera
35	593	15.3	550	1	A34576	crystal protein pr
36	588	15.2	767	2	S47639	acetylcholinestera
37	581.5	15.0	746	2	A25363	acetylcholinestera
38	575	14.9	540	2	S51043	carboxylesterase (
39	572.5	14.8	602	2	T37254	acetylcholinestera
40	570.5	14.7	637	2	S66236	acetylcholinestera
41	569.5	14.7	554	2	T31783	hypothetical prote
42	569.5	14.7	557	2	A56690	esterase Caenorh
43	562.5	14.5	545	2	A89046	protein B0238.1 11
44	557.5	14.4	593	1	S25062	triacylglycerol 11
45	553	14.3	956	2	A56920	gliotactin precurs

ALIGNMENTS

RESULT 1
A47162
Cholinesterase B (BC 3. -.-) precursor - mallard
C/Species: Anas platyrhynchos (mallard)
C/Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004
C/Accession: A47162
R./Hwang, C.S.; Kolatukudy, P.E.
J. Biol. Chem. 268, 14278-14284, 1993
A./Title: Molecular cloning and sequencing of cholinesterase B cDNA and stimulation of expe
A./Reference number: A47162; MUID:93300823; PMID:8314791
A./Accession: A47162
A./Status: preliminary
A./Molecule type: mRNA
A./Residues: 1-557 <HMA>
A./Cross-references: UNIPROT:Q04791; GB:L05493; NID:G213100; PIDN:AAA49223.1; PID:G213101
C./Superfamily: cholinesterase; cholinesterase homology
C./Keywords: hydrolase
F./56-545/Domain: cholinesterase homology <CHE>

Alignment Scores:

Pred. No.: 7,1e-91 Length: 557
Score: 1298.50 Matches: 265
Percent Similarity: 62.77% Conservative: 79
Best Local Similarity: 48.36% Mismatches: 177
Query Match: 33.54% Indels: 27
DB: 2 Gaps: 8

US-10-023-515-1 (1-2158) x A47162 (1-557)
QY 183 ACTGGGCTTGTGTTGAAGGGCCACAGAGAACACAGCTGGATGATTCAGGGCAAG 242
DB 23 ThrGlyGlnValAlaGlnGlnProGlnValValThrAenTYrGlySerValArgIlyr 42
QY 243 CAAGTCATGTCGTGGGAAGCCCTGCTGTGAACGTTCTCTGGAGTCCCTTTGCT 302
DB 43 GlnValIyValAlaenAlaGlnValArgSerValenValPheLeuGlyLeuProheAla 62
QY 303 GCTCCCGCGTGGATCCCTGATTTACGAACCGCAGCGCTGATCCCTGGATAC 362

```

Db      63  LysProValIglyProLeuPhgSerGluProGlnProProGlnProTriblyGly 82
Qy      363  TTGGAGAAAGCACTCTTCACTTAATTTGGCTCCGAACTCAGAGTGGCTCTTCA 422
Db      83  ValAlaGAPhAlaAlaSerTyProProMetCysLeuGlnMetPys-----ValLeu 99
Qy      423  GATCAACATGCTC-----AAGGTGATTAACCGGAATTCGAGGTGCA 467
Db      100  GlyGlnTyLeuSerAAspAlaIleThrAsnArglySgIuLysValAlaGluGlnIleSer 119
Qy      468  GAAAGTCTCTTCACTTAACAATCTATGCGCTCCCAAGGATACAGGCTCCAGCTC 527
Db      120  GlnuAPcysLeuTyLeuAsnValTyThrProValSerThrGlnGlnGlnGlnGln 139
Qy      528  CCCGCTTGGTGGTGGTCCGAGAGGAGGCTTCAAGAGTGGCTCAGCTCTTGGAT 587
Db      140  ProValPheValIleThrIleSgIglyGlyLeuValSerGlyAlaAlaSerSerTyAsp 159
Qy      588  GAGTCCCGCTGAGCTGCTATGAGACATGCTGGTGGTGGTCCAGTCCGAGTACGA 647
Db      160  GlySerAlaLeuAlaAlaPheAAspAsnValValValAlaThrIleGlnTyAlaGly 179
Qy      648  AATTTGGTCTTCCACCAATGAGATCAGATGCTCCGGGAACTGGCTTCAAGAC 707
Db      180  IleAlaGlyTyThrPheSerThrGlyAspLysAlaAlaArgIlyAsnTrpGlyTyLeuAsp 199
Qy      708  CAGTGGCTCTGCTGCTGCTCCAGAAACATGCTGCTGGTGGGAGCCGAC 767
Db      200  GlnValAlaAlaLeuGlnTrpIleGlnGlnIleAlaIleAlaPheArgIlyAspProGly 219
Qy      768  TCTGTACCATCTTTGGCGAGTCCGGGAGCAATAAGTGTCTTATGTTACTGTCT 827
Db      220  SerValThrIlePheGlyGlnSerAlaGlyValSerValSerAlaLeuValLeuSer 239
Qy      828  CCCATGCGCAAGGCTTATTCACAAAGCATCATGAGAGGGGTGGCCATC---ATC 884
Db      240  ProLeuAlaLysGlyLeuPheHisIleValAlaIleSerGlnSerGlyThrAlaValArgIle 259
Qy      885  CCTTACCTGAGGCCCATATTATGAGAAAGTGAAGCATGCTGAGTGGTTCACATTTC 944
Db      260  LeuPheThrGlu-----GlnProGlnGlnGlnAlaGlnAlaArgIleAlaAla 275
Qy      945  TGGGTAAACATGCGTCACTCATGAGGCTCTGAGTGGCTTCAAGCAAAACCTTCC 1004
Db      276  AlaGlyCysGlnLysSerSerSerAlaAlaLeuValGlnCysLeuAlaGlnTyThrGlu 295
Qy      1005  AAGGAGCTG-----CTGACCTTCAAGCCAGAAACAAAGTCTTTCATCTGAGTGGTGTAT 1058
Db      296  AlaGlnMetGlnGlnIleThrLeuLysMetProPheMetPheIleSerAlaSerLeuAsp 315
Qy      1059  GGTGCTTCTTCTTCAATGAGCTCTAATCTATTGCTTCAAGAAAGCATTTAAAGCAATT 1118
Db      316  GlyValPhePheProLysSerProArgGlnLeuLeuSerGlyLysValIleAsnAlaVal 335
Qy      1119  CTTTCCATCATCGAGTCAATTAACACAGAGTGGCTTCCGCTGCCCT---ATGAGAG 1175
Db      336  ProTyIleIleGlyValAsnAsnCysGlnPheGlyTrpIleLeuProArgMetMetLys 355
Qy      1176  GCTCTGAGATCTCTGAGTGGCTCCAAAGTCCCTTGGCTCATCTGATCAAAACATC 1235
Db      356  PheProGlnPheThrGlnGlyLeuGlnLysAspValAlaArgGlnValIleGlnIleSerThr 375
Qy      1236  CTGCAACATC-----CCGCTCAAGTATTGGACCTTTGGGCTTAATGAATAC 1280
Db      376  LeuAlaLeuSerPheLysGlyAlaProSerAspIleValAspLeuValTyAsnGlnTyLys 395
Qy      1281  TTCCATGACAAAGCATCCCTGATGAAATCCGAGACAGTCTTGGACTTGGAGAT 1340
Db      396  IleGlyValAlaGlnAsnAlaGlnValAlaArgAspGlyLeuLeuAspSerIleAlaAsp 415
Qy      1341  GTGTTCTTTGGTGGTCCCTGACATGATACAGCTGATATACAGAGATGCTGTGACCT 1400

```

```

Db      416  ProLeuPheValPheSerAlaValAlaAlaArgHisIleArgAspAlaGlyAsnPro 435
Qy      1401  GTCTACTTTCATGAGTTTCGACCGGCTCAGTCTTTCGAAAGACACAGACCGGCTTTT 1460
Db      436  ValTyThrPheTyGlnPheGlnHisArgProSerSerAlaAlaGlyValValProGlnPhe 455
Qy      1461  GTCAAGCCGACACGCTGATGAAAGTCCGCTTGTTCGGTGGTGGCTTCTGAAAGGG 1520
Db      456  ValLysAlaAspHisAlaAspGlnIleAlaPheValPheGlyLysProPheLeuAlaGly 475
Qy      1521  GACATTGTTATTTGGAAGAGCCACGAGAGAGAAATTACTGAGCCGGAAGATGATG 1580
Db      476  Asn-----AlaThrGlnGlnGlnAlaLysLeuSerArgThrValMet 489
Qy      1581  AATATCGGGCTACCTTTGCTGAAACCGGAATCTTAATGGGAACGACTGTCTGTGG 1640
Db      490  LysTyTrpThrAsnPheAlaAlaArgAsnLysAsnProAsnGlnGlnGlyLeuValHisTrp 509
Qy      1641  CGAGTTTAAATCTGACATGAGACATCTCCAGCTGACATTGAACATGAGCTTCGACAG 1700
Db      510  ProGlnTyAspMetAspArgLysArgTyLeuGlnIleAspLeuThrGlnLysAlaAlaLys 529
Qy      1701  AGACTCAAGAAACCGGGGTGATTTTGAACAGACATCCCGCTGATCTGTCTGCC 1760
Db      530  LysLeuLysGlnAlaGlyLysMetGlnPheTrpMetGlnLeuThrGlnGlnIleMetSerAsp 549
Qy      1761  TCCGACATGCTCCACAGTCTCTT 1784
Db      550  ArgArgArgLysHisIleThrAspLeu 557

RESULT 2
A34329
60K esterase (BC 3.1.1.-) isoform 2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004
R:Accession: A34329
R:Ozols, J.
J. Biol. Chem. 264, 12533-12545, 1989
A:Title: Isolation, properties, and the complete amino acid sequence of a second form of
A:Reference number: A34329; MUID:89308686; PMID:2745458
A:Accession: A34329
A:Status: preliminary
A:Statue: preliminary
A:Molecule type: protein
A:Residues: 1-532 <OZO>
A:Cross-references: UNIPROT:P14943
C:Superfamily: cholinesterase; cholinesterase homolog
C:Keywords: carboxylic ester hydrolase
F:32-517/Domain: cholinesterase homolog <CHE>
F:201,430/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 9,13e-86 Length: 532
Score: 1231.50 Matches: 253
Percent Similarity: 62.19% Conservative: 76
Best Local Similarity: 47.83% Mismatches: 167
Query Match: 31.81% Indels: 33
DB: 2 Gaps: 7

US-10-023-515-1 (1-2158) x A34329 (1-532)
Qy      204  CCAACAGAGAAACAAGCGTGGATGATTCAGGCAAGCAAGTCACTGTCTGGAGAC 263
Db      6  ProIleArgAsnThrHisThrGlyValAlaArgLysSerLeuValHisValGlnGlyThr 25
Qy      264  CCTGTGCGCTGAGAGTCTCTGGAATCCCTTTGGTGGTCTCCCGCTGGATTCCTG 323
Db      26  AspAlaGlyValHisThrPheLeuGlyIleProPheAlaLysProProLeuGlyProLeu 45
Qy      324  CGATTTAAGAAACCGCAAGCTGCAATCGCTGGAGTAACTTGGAGAAACCAAGCTTCAC 383
Db      46  ArgPheAlaProProGlnProAlaGlnAlaTrpSerGlyValAlaGAspGlyThrSerLeu 65
Qy      384  CCTAATTGTGCTCCAGAACTCAGAGTGGCTGCTTGAATCAACATGCTCAAGGTG 443

```

Db	66	ProAlaMetCvbleuGlnAhn-----LeuAlaIleMetAerGlnAerValIleuLeuLeu	83
Qy	444	CATTAC-----CCGAAATTCGAGTGTCAAGAAGATCCTCTACCTGAACATCTATGCG	497
Db	84	HisPheThrProProSerIleProMetSerGluAerCvbleuYtyleuAhnIleYtSer	103
Qy	498	CTGCGCCACCGCGGATACAGGCTCCAAAGTCCCCCGCTTGCTGGTGGTCCCAAGAGTCC	557
Db	104	ProAlaHisValAaerGluGlySerAerProValMetValThrIleHisGlyYcYcIy	123
Qy	558	TTCAAGATCGAGTCAAGCTCATCTTTATGGTGTCCGCGCTGGCGCTATAGAGAGATG	617
Db	124	LeuThrMetGlyMetAerAerSerMetYtAerGlySerAlaIleAaIaAerPheGluAerAl	143
Qy	618	CTGGTGTGTGTGTCAGTACCTCCGCTAGAAATATTTGGTTCTTCCACACATGGAGTAC	677
Db	144	ValValValAlaThrIleGlnTyrArgLeuGlyValLeuGlyPhePheSerThrGlyAerProI	163
Qy	678	CATGCTCCGGGAACTGGGCGCTTCACAGACCAAGATGGTGTCTGTCTCTGGAGTCCAGAG	737
Db	164	HisAlaThrGlyAerHisGlyTyrTyleuAerGlnValAlaAlaIleAaXGTrValGlnIly	183
Qy	738	AACATCGAGTCTTCGGTGGGAGCCCAAGCTCTGTACACATCTTTGGGAGTCCGGGGA	797
Db	184	AerHisIleAaIaAerPheGlyGlyAerProGlyAaerValAlaThrIlePheGlyHiserAlaGly	203
Qy	798	GCCATTAAGTCTTCAAGTCTTATATGTCTCCCATGGCCAAAGGCTTATTCACAAAGCC	857
Db	204	GlyThrSerValSerSerHisValIleAerProMetSerGlnGlyLeuPheHisGlyYala	223
Qy	858	ATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGGCCCATGATTAAGAAAGT	917
Db	224	IleMetGlnSerTyleuValAlaIleuLeuProGlyLeuIleThrSerSerGlnValVal	243
Qy	918	GAGAGCTCGACGTGTTCACATTTCTGTGGTAAACATGGGTCAAGTCTCGAGGCTGT	977
Db	244	Ser-----ThrValValAlaAerHiserAerGlyGlyGlnValAaerSerGlnThrIleu	261
Qy	978	CTGAGGTGCTCGAGGACAAACCCCTCCAGAGCTGAGCCCTCAGCCAGAAACAAAG	1037
Db	262	ValArgCvbleuAaAerGlyAaIySerGlnGluGluMetSerAlaIleThrGlnValAerMet	281
Qy	1038	TCTTTCATCGAGTGTGTATGTGTCTTTCTTCTTAATGAGCTCTAGATCTATGTCT	1097
Db	282	LeuIleProGlyValValAerProIyAaIaPheLeuProAerHisAerProGlnGluLeuAaIa	301
Qy	1098	CAGAAAGCATTTAAAGCAATCTCTTCCATCATCGGAGTCAATTAACAGAGTGGCTTC	1157
Db	302	LeuAlaAerPheGlnProValProSerIleIleGlyIleAaAaAerGlyTyrGlyTyr	321
Qy	1158	CTGCGCGCT-----ATGAGGAGGCT	1178
Db	322	IleIleProIyAerLeuLeuLeuAlaIleAerProGlnGluAaIyAaerAerGlnAaIaMet	341
Qy	1179	CTGAGATCTCTCAGTGGCTCCACAAAGTCCCTTGCGCTCATGTATCAAAACATCCTG	1238
Db	342	ArgGluIleMetHisGlnAlaIleThrIleGlnIleuMetIleu-----	354
Qy	1239	CACATCCCGGCTCAGATATTGCACTTGTGGCTAATGAATATCTTCCATGACAGAGCTCC	1298
Db	355	-----ProProAlaIleuGlyAaerPheLeuMetAerGlnTyrMetGlySerAerGlnIyAer	372
Qy	1299	CTGACTGAATATCCGAGACAGTCTTGTGACTCTGCTGAGAAATGTGTCTTTGTGGTCCCT	1358
Db	373	ProIyHisIleuMetAaIaGlnPheGlnIleIleuMetAaAaerAlaMetPheValMetSerPro	392
Qy	1359	GCAGTGTATCAGAGTGCATATCACAAGATGCTGGTGTCAAGCTGTACTTATAGAGTTT	1418
Db	393	AlaIleuAaerGlyAlaIleAerLeuGlnAaIyG---SerHisAaIaerProIyAerPheTyrGluPhe	411
Qy	1419	CGGACCGGGCTCAGTGTCTTGAAGACAGAACCGGCTTTGTCAAGCCGACAGAGCT	1478

Db	412	glnhivarkproseerhethilyadavleuakgrvotrhniavlaadavlaadhrni:6Ly	431
Qy	1479	GATGAAGTCCGCGCTTTGGTTCGAGTGGCGCTCTCTGAAGGGGACATGTTGATGTTGGAA	1538
Db	432	AapiluvaiValaPhevalPhealArgSerhivleuPheGlySerilyValaProlen-----	449
Qy	1539	GGAGCCACGGAGAGAGAAATTACTAGAGCCGGAAGATGATGAATATCTGGCTGACCTTT	1598
Db	450	-----ThrgluGluGluGluLeuLeuSerhlyglValMetlyGlyTyrPralaaPhe	467
Qy	1599	GCTTGAACCGGGAATCTTAATGGGAAGACSTGTCGTCTGGTGGCCAGCTTAATCTGAAT	1658
Db	468	AlatrgsanharganProlenGlyGluGlyLeuAlaIhivTyrProlenPheadervleuavr	487
Qy	1659	GAGCAGTACCTTCAGCTGGACTTGAACATGAGCTCGGAGCAGACATCAAGACCCGGG	1718
Db	488	GlnArgTyrLeuGluGlnLeuAsnMetGlnProlValGluGlnAlaIatLeuLyvalaaArgx	507
Qy	1719	GTCGATTTTGGACAGACCATCCGCC	1745
Db	508	LeuGlnPheTyrPrinhIsthrLeuPro 516	

```

RESULT 3
JC5408
carboxylesterase (EC 3.1.1.1) - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: JC5408
R:Schwer, H.; Langmann, T.; Daig, R.; Becker, A.; Aslaniadis, C.; Schmitz, G.
Biochem. Biophys. Res. Commun. 233, 117-120, 1997
A:Title: Molecular cloning and characterization of a novel putative carboxylesterase, pTCE
A:Reference number: JC5408; MUID:97289502; PMID:9144407
A:Accession: JC5408
A:Molecule type: mRNA
A:Residues: 1-559 <SCH>
A:Cross-references: UNIPROT:000748; GB:Y09616; NID:g2058317; PIDN:CAA70831.1; PID:g2058317
A:Experimental source: intestine
C:Comment: This enzyme hydrolyzes many xenobiotics, such as carboxyl esters, thioesters &
C:Genetics:
A:Gene: GDB:CE82; ICE; CE2
A:Cross-references: GDB:9959011
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:58-544/Domain: cholinesterase homology <CH>
F:15-95,123-280,291-428/Disulfide bonds: #status predicted
F:111,276/Binding site: carbohydrate (asn) (covalent) #status predicted
F:228,457/active site: Ser, His #status predicted

Alignment Scores:
Pred. NO.: 1,21e-85 Length: 559
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 31.77% Indels: 24
DB: Gaps: 8

US-10-023-515-1 (1-2158) x JC5408 (1-559)

QY GGGCCTTGCTGGAAGGGCCACAGAGCAACACCAGGCTGGATGATTCAGGCGCAAGCA 245
Db ::::: ||||| :::::
26 GlylnapSerAlaSerProIleArgThrThrHisstHglylValInvalIuGlySerLeu 45
246 GTCACTGATCGGGAAGCCCTGTGCTGTGCAACGATTCCTCGAGAGCCCTTGCTGTCT 305
Db ||||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
46 ValHisValIysGlyAlaAsnAlaGlyValGlnThrPheLeuGlyIleProPheAlaIys 65
306 CCCCCGCTGGGATCCCTCGATTTCAGAACCCGCAAGCCTGCATCGCCTGGGATTAATTG 365
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
66 ProProIuGlyProLeuArgPheAlaProProGluProProGluSerIrpSerGlyVal 85
366 CGAAGACCACTCTACCCCTTAATTGTGGCTCCCAAGAC-----TCAGAG 410
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
86 ArgAspGlyThrThrHisProAlaMetCysLeuGlnAspLeuThrAlaValGluSerGlu 105

```

```

QY 411 TGGCTGCTTTAGATCAACATGCTCAAGTGCATTACCCGGAATTCCAGTGTCAAG 470
    |||
Db 106 PheLeu-----SerGlnPheAbnMetThrPheProSerAspSerMetSerGln 121
QY 471 GACTGCTCTTACCTGAACATCTATGCGCTGCCACGCCGATAGAGGCTCCAGCTCCC 530
    |||
Db 122 AspCysLeuTyrLeuSerIleTyrThrProAlaHisSerHisGluGlySerAbnLeuPro 141
QY 531 GCTTGGAGTGGATCCAGAGAGTGGCTTCAAGACTGGCTCAGCTTCATTTTGATGG 590
    |||
Db 142 ValMetValIlePheHisGlyAlaLeuValPheGlyMetAlaSerLeuTyrAspGly 161
QY 591 TCCGCTGCTGCTGCTTGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
    |||
Db 162 SerMetLeuAlaIleLeuGluAbnValValValIleIleGlnTyrAbnLeuGlyVal 181
QY 651 TTTGGTTTCTTACACACATGGAGATCAGATGCTCCGGGAACTGGGCTTCAAGACCA 710
    |||
Db 182 LeuGlyPhePheSerThrGlyAspIleHisIleThrGlyAbnTyrGlyLeuAspGln 201
QY 711 GTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
    |||
Db 202 ValAlaIleLeuAlaGlyThrValGlnGlnAbnIleAlaHisPheGlyGlyAbnProAspArg 221
QY 771 GTGACCATCTTTGGCGAGTCCGCGGAGCCATAAGTCTTCTAGCTTATAGTCTGCC 830
    |||
Db 222 ValThrIlePheGlyGlySerIleAlaGlyIleThrSerAlaSerSerLeuValValSerPro 241
QY 831 ATGGCCCAAGGCTTATCCCAACAAAGCATGAGAGTGGGCTGCTCATCATCTTAC 890
    |||
Db 242 IleSerGlnIleLeuPheHisGlyAlaIleMetGlnSerIleValAlaLeuLeuProGly 261
QY 891 CTGAGAGCCCATGATTATGAGAAAGTGAAGACTG-----CAGGTGGTGAACATTTC 944
    |||
Db 262 LeuIleAla-----SerSerAlaAspValIleSerThrValAlaAbnLeu 277
QY 945 TGTGGTAAACATGCGTCAGACTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
    |||
Db 278 SerAlaCysAbnGlnValAbnSerGlnAlaLeuValGlyCysLeuAbnGlyIleSerLeu 297
QY 1005 AAGGAGCTGCTGACCTCAGACCCCAAAACAAAGTCTTTCACGCTGCTGCTGCTGCT 1064
    |||
Db 298 GlnGluIleLeuAlaIleAbnHisProPheIleMetIleProGlyValAlaAbnGlyVal 317
QY 1065 TTCTTTCTTAATGAGCTCTAAGATCTATGCTGCTCAAAAGATTAAAGCATTCCTCC 1124
    |||
Db 318 PheLeuProArgHisProGlnIleLeuLeuAlaSerAlaAspPheGlnProValProSer 337
QY 1125 ATCATCGAGTCAATTAACACAGAGTGTGCTTCTGCTGCTGCT-----ATGAAG 1172
    |||
Db 338 IleValGlyValAbnHisAbnGlnIlePheGlyTyrPheIleProIleValAlaMetArgIleTyr 357
QY 1173 GAGGCTCTGAGATCCTCAGTGGCTCAACAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
    |||
Db 358 AspThrGlnIleGlyMetAspArgGlnIleAspGlnAlaIleLeuGlnIleMetLeuThr 377
QY 1233 ATCCGCAACATCCGCTCAGATTTGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1292
    |||
Db 378 LeuLeuMetLeuProThrPheGlyAspLeuLeuArgGlnIleTyrIleGlyAspAbn 397
QY 1293 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352
    |||
Db 398 GlyAspProGlnIleThrLeuGlnAlaGlnIlePheGlnIleMetMetAlaAspSerMetPheVal 417
QY 1353 GTCCCTGCACTGATACAGCTGCTGATATACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 1412
    |||
Db 418 IleProAlaLeuGlnValAlaIleHisPhe---GlnCysSerArgAlaProValIleTyrPheTyr 436
QY 1413 GAGTTTGGCAACCGGCTCAGTGGCTTGAAGACACAGAACCGGCTTGTTCACAAAGCCGAC 1472
    |||
Db 437 GluPheGlnHisGlnProSerTyrPheLeuYenAbnIleArgProProHisMetLeuValAlaAsp 456

```

```

QY 1473 CAGCTGATGAAGTCCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1532
    |||
Db 457 HisGlyAspGluLeuProPheValPhe---ArgSerPhePheGlyGlyAbnTyrIleLeu 475
QY 1533 TTGGAAGAGCCACGAGAGAGAGAAATTACTGACCCGGAAGATGATGAATTAATCTGGCT 1592
    |||
Db 476 Phe-----ThrGlnIleGlnIleGlnIleLeuSerArgIleMetMetLeuTyrTyrAla 492
QY 1593 ACCTTGCTGCAACCGGGAATCTTAATGGAACGACCTGCTGCTGCTGCTGCTGCTGCTGCT 1652
    |||
Db 493 AsnPheAlaArgAbnGlnAbnProAbnHisGlnIleGlyLeuProHisTyrPheLeuPheAsp 512
QY 1653 CTGACTGAGCACTTACTCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1712
    |||
Db 513 GlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 532
QY 1713 CCGCGGTGATTTTGGACGACGACCATCTCCC 1745
    |||
Db 533 HisArgLeuGlnPheThrIlePheValAlaLeuPro 543

```

```

RESULT 4
S47655
carboxylesterase (EC 3.1.1.1) precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S47655
R:Some, T.; Isode, M.; Takabatake, E.; Wang, C.Y.
Biochim. Biophys. Acta 1207, 138-142, 1994
A>Title: Cloning and sequence analysis of a hamster liver cDNA encoding a novel putative
A:Reference number: S47655; MUID:94318665; PMID:8043605
A:Accession: S47655
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <SD>
A:Cross-references: UNIPROT:Q64419; EMBL:D28566; NID:9531238; PIDD:BA05913.1; PIDD:95312;
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
P:58-546/Domain: cholinesterase homology <CHI>
P:227,459/Active site: Ser, His #status predicted

```

```

Alignment Scores:
Pred. No.: 2.8e-83 Length: 561
Score: 1199.00 Matches: 252
Percent Similarity: 61.16% Conservative: 74
Best Local Similarity: 47.28% Mismatches: 181
Query Match: 30.97% Indels: 26
DB: 2 Gaps: 7

```

```

US-10-023-515-1 (1-2158) x S47655 (1-561)
QY 186 GGGCTTCTGCTGAAGGCGCACAGAGAACACAGGCTGGATGGATTTCAGGGCAAGCA 245
    |||
Db 26 GlyGlnAspSerValSerProIleArgAbnHisThrGlyGlnValAlaArgIleLeu 45
QY 246 GTCACTGCTCTGGGAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
    |||
Db 46 ValTyrValIleGlyGlnIleValAlaThrGlyValIleAlaPheLeuGlyIleProPheAlaLeu 65
QY 306 CCCCCGCTGGATCCCTGCGATTTACGAACCGGACCTGCATCGCCCTGGAGTAATCTTG 365
    |||
Db 66 ProIleValGlyProLeuArgPheAlaProProGluProProGluProTyrSerGlyIle 85
QY 366 CGAAGACCACTCTTACCTTAATTTGGCTCCGAGACTCAGAGTGGCTGCTGCTGCTGCTGCT 425
    |||
Db 86 ArgAspGlyThrSerGluProAlaMetCysLeuGlnIleThrAspPheMetArgProGlnIle 105
QY 426 CAACACATGCTTAAGCTGCATTAACCGAAATTCGAGTGTCAAGAGTGCCTTACCTG 485
    |||
Db 106 SerIleGluArgGlyIleIleLeuProThrIleSerMetSerGlnAspCysLeuTyrLeu 125
QY 486 AACATCTATGGCGCTGCGACGCGATACAGGCTCCCAAGCTCCCGCTTGTGTGCTTC 545
    |||
Db 126 AsnIleTyrThrProAlaHisIleGlnIleGlySerAbnLeuProValMetValTyrIle 145

```


QY	546	CAAGAGGCTGCTTCAAACTGCGTCAACCTTCACCTTGTATGGCTCCGCGCTGCTGAC	605
Db	146	HtSgIyglYalAlaLeuValMetCylIleMetAlaSerMetAlaNapGlySerLeuLeuAlAla	165
QY	606	TATGAGACGCTGCTGTGTGTGCTGTCCAGATCCGGCTAGAGAAATATTTGGTTCTTCACC	665
Db	166	ThrgIuAepIleValIleValSerIleGlnTyrzhrLeuGlyIleLeuGlyPhePheSer	185
QY	666	ACATGGGATCAGCATGCTTCGCGGGAACCTGGGCTTCAAGACACAGTGGCTGTCTGCC	725
Db	186	ThrgIyAepGluHtAlaIArgGlyAsnTrpGlyTyrLeuAepGlnValAlAlaLeuHtS	205
QY	726	TGGGTCGAGAAACATCGAGTCTTTCGGTGGGGACCCGACCTGTGATCATCTTTGGC	785
Db	206	TrpValGlnGlnAsnHtAlaSerPheGlyIleGlyAsnTrpGlyGlnValHtHtIlePheGly	225
QY	786	GAGTCCGCGGAGCCATAGTGTCTTAGTCTTATCTGTCTCCATGCGCAAGGCTTA	845
Db	226	ValSerAlaGlyGlyThrSerValSerSerLeuValValSerPrometSerGlyLeu	245
QY	846	TTCCACAAAGCATCATGATGAGAGTGGGGTGCACATCCCTTACCTTCGAGAGCCCATGAT	905
Db	246	PheHtSgIyAlAlIleMetGlnSerCylValAlaLeuLeuProAepLeu--IleSerAsp	264
QY	906	TATGACAAAGTGAAGACCTCGAGGTGTGCACATTTCTGTGGTAACAATGCGTCAGAC	965
Db	265	ThrProGluAlaValTyrThrProValValAlaAsnGlnSerGlyCySgIuAlbAAsp	284
QY	966	TCTGAGCCCTGCTGAGGTGCTCTGAGACAAACCTTCAGAGAGCTGTACCTCAGC	1025
Db	285	SerGluAlaLeuValHtSLeCylAeuAArgIuTyrThrgIuAlaGluIleLeuAlAlIleAsn	304
QY	1026	CAGAAACAAAGTCTTTCACTCGAGGTGTGATGGTCTTTCTTCCATTAAGCGCTCA	1085
Db	305	GlnValPheIleMetThrProGlyValValAspGlyIlePheLeuProAArgHtIleProGln	324
QY	1086	GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTCTTCATCATCGAGTCATTAACAC	1145
Db	325	GluLeuLeuAlaSerValAspPheHtIleProValProSerIleIleGlyValAspSerAsp	344
QY	1146	GAGTGGCTTCCTGCTGCTCATG-----AAGAGGCT	1178
Db	345	GlnCySgIyTrpGlyValAlaProLeuHtMetGlyLeuAepHtValIleLeuAsnHtIleThr	364
QY	1179	CCTGAGATCTCATGAGTGTCCACAAACATGCTGCTGCTCATCTGATCAAAACATCTCTG	1238
Db	365	ArgGluThrLeuProAlaPheLeuIuYSerTrpAlaGluHtIleMetMet-----	380
QY	1239	CACATCCCGCTCAGTATTTGACCTTGTGGCTAATGAATCTTCATGACAGACATCC	1298
Db	381	--LeuProProGlnCySLeAspPheLeuLeuMetGlnGlyTyrMetGlyAspValAlGluAsp	399
QY	1299	CTGACTGAATATCCGAGACATGCTTCTGACATCTTGTGAGAAATGNTTTTGTGGTCTCT	1358
Db	400	ProGlnThrLeuGlnAlaGlnPheAArgIuLeuMetCyAspPheMetPheValIlePro	419
QY	1359	GCATGTATCAGCTCATATCACAGAGATCTGGTGCACCTGTCTACTTATGAAGTT	1418
Db	420	AlaLeuIuYValAlaIatYrPheGluHtArg--SerHtAlaProValTyrPheTyGluPhe	438
QY	1419	CGGACCGGCTCAGTCTTT-----GAAACACAAAGCGGCTTTGTGCAAGCCGAC	1472
Db	439	GlnHtSgIuSerSerPheIleLeuAsnIuYAspAlaArgProSerHtAlaArgAlaAsp	458
QY	1473	CACGCTATGAAGTCCGCTTTGTGTCTGCGTGGTCTCTC-----CTGAAAGGGAGACAT	1526
Db	459	HtSgIyAepHtAlaAlaPheValPheGlySerAspPheTrpGlyLeuIuYIleAspLeu	478
QY	1527	GTTATGTTCAAGAGCACCGAGAGAGAAAGTTATCTGACCCGGAAGATGATGAATAC	1586
Db	479	-----ThrgIuGlnIuYLeuLeuAsnIuYAspGlyMetMetIleTyr	492

```

QY      1587  TGGGCTACCTTTGCTCGAACCCGGGATCTCTAAATGGGAAGACACCTGTCTGTGGCAGCT 1646
      |||||
Db      493   TTPAAlasnpPhalAlaRghi sglYlanpProasSercIuclYleuProIyrtfPProIu 512
      |||||
QY      1647  TATAATCTGACTGAGCAGCATCTCCAGCTGACTGACTGAAATCATGAGCCTCGACAGAGACTC 1706
      ::|||
Db      513   LeuValhIsAspAlaspIIntYrleuYlLeuAspIleclnProAlaValGlyArgAlaleu 532
      ::|||
QY      1707  AAAGAACCGCGGGGTGATTTTGTGACACGACCATCCCC 1745
      |||
Db      533   LysSerArgYlLeuHIspeHtPTrhYlLeuPro 545
      ::|||

RESULT 5
S34607
carboxylesterase (BC 3.1.1.1) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S34607
R:Aida, K.; Moore, R.; Negishi, M.
Biochim. Biophys. Acta 1174, 72-74, 1993
A:Title: Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase in
A:Reference number: S34607; MUID:93326638; PMID:7916639
A:Accession: S34607
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <Alt>
A:Cross-references: UNIPROT:063880; GB:S64130; NID:G404388; PIDN:AA827606.1; PID:G404389
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:46-536/Domain: cholinesterase homology <CHR>
I:215,443/Active site: Ser, His #status predicted

```

Alignment Scores:		
Pred. No.:	2,526-82	Length: 554
Score:	1186.50	Matches: 256
Percent Similarity:	61.16%	Conservative: 81
Best Local Similarity:	46.46%	Mismatches: 183
Query Match:	30.64%	Indels: 31
DB:	1	Gaps: 10

US-10-023-515-1 (1-2158) x S34607 (1-554)

QY	132	TGCTTTTCTGATTTCTCCAGCCCTGTTGGGACACAGACAGTGGGAAAAATGGGCTT	191
Db	3	CysLeuLeuLeuIlePheProThrThrVal-----IleGlyPro	15
QY	192	TCTGTGAAGGCGCCACAGAGAACACACAGCTGGGATGATTACAGGCGAACAGTCACT	251
Db	16	LybValThrGlnProGlnValAspThrProLeuGlyArgValArgGlnValGly	35
QY	252	GTGCTGGGAAGCCCTGTGCTGTGAAACGTTGCTCTGGAGTCCCTTTGGTGTCTCCCG	311
Db	36	ValLybAspThrAspArgMetValAspValPheLeuGlyIleProPheIleGlnIlePro	55
QY	312	CTGGGATCCCTGGCGATTTACGAACCCGACGCTGCATCGCCCTGGGATTACTTGGAGA	371
Db	56	LeuGlyProLeuArgPheSerIleProLeuProProGlnProIleProGlnGlyValArgAsp	75
QY	372	GCCACCTCCATCCCTTAATTTGTGCTTCCAGAACTACAGAGTGGGCTCTTAAGATCAAC	431
Db	76	AlaSerIleAspProPheMetCysLeuGlnAspValGlnArgMetSerAspSerArgPhe	95
QY	432	ATGCTCAGAGTGCATTACCCGAAATTCCGAGTGTCCAGAAAGTCCGCTACCTGAACATC	491
Db	96	ThrLeuAsnGlnLysMetLysIlePheProIleSerGlnAspCysLeuThrLeuAsnIle	115
QY	492	TATGCGCTGCCACGCGCATACAGGCTCCAAAGTCCCGCTTGTGTGTGTTCCSAGA	551
Db	116	TyrSerProThrGlnIleThrIleGlnArgLysArgProIleMetValTrpIleIleGly	135
QY	552	GATGCTTCAACAGCTGGCTCAGGCTCCATCTTGAATGGGTCCGGCTGGGCTGCTAAGAG	611
Db	136	GlySerIleArgSerValGlySerSerThrSerHisAspGlySerIleAlaMetValIleValGly	155


```

Qy 675 CAGCATGCTCCGGGGAACCTGGGCTTCAAGACGAGTGGCTGTCTGTGCTGGTCCG 734
Db 183 GlnHisSerProGlyAsnThrPalaHisLeuMetGlnLeuAlaHisLeuArgTrpValGln 202
Qy 735 AAGAACATCGAGTTCTTGGGTGGGAGCCCAAGCTGTGTGACCATCTTGGCCGAGTCCG 794
Db 203 AsnProteinAlaAsnPheGlyGlyLeuAsnProAspSerValThrIlePheGlyGlyLeuSer 222
Qy 795 GGAGCCATTAAGTGTCTTACTGTCTTACTGTCTTCCATGGCCAAAGGCTTATTCACAA 854
Db 223 GlyGlyIleSerValSerValLeuValLeuSerProLeuGlyValAspLeuPheHisArg 242
Qy 855 GCCATCATGAGAGGAGGTGGGCGCCATCCCTTACCGAGAGCCCATGATATGAGAG 914
Db 243 AlaIleSerGlnSerGlyValValIleAsnThrAsnValGlyValLeuAsnIleGlnAla 262
Qy 915 AGTGAGACCTGACAGTGGTGGACATTTCTGTGTAAACATGCGTCAAGACTGTAGGCC 974
Db 263 ValAsnGlnIleIleAlaThrLeuSerGlnCys-----AsnAspThrSerSerAlaAla 280
Qy 975 CTGCTGAGGTGGCTGTGAGACAAACCTTCAGAGGCTGTGACCTCAGC----- 1025
Db 281 MetValGlnCysLeuArgGlnIleThrGlnSerGlnLeuLeuGlnIleSerGlyValLeu 300
Qy 1026 ---CAGAAACAAAGCTCTTACCTGAGAGTGTGTGGTCTTCTTCTTATGAGCCT 1082
Db 301 ValGlnIleAsnIleSerLeuSerThrMetIleAspGlyValValLeuProValAspPro 320
Qy 1083 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCATTCCTTCCATGATCGAGTCAATAAC 1142
Db 321 GlnGlnIleLeuAlaGlnIleSerPheAsnThrValProIleValGlyPheAsnIle 340
Qy 1143 CAGGAGTGTGGCTCTGCTGCTGCTATG-----AAGAGGCTCTGAG---ATCTTC 1190
Db 341 GlnGlnPheGlyTrpIleIleProMetMetLeuGlnAsnLeuLeuProGlnIleGlyMet 360
Qy 1191 AGTGCTCCAAACAGTCCCTGCTCCATCGATACAAACATCCCTGACATCCCGCT 1250
Db 361 AsnGlnIleThrAlaSerLeuLeuLeuArgArgPheHisSerGlnLeuAsnIleSerGln 380
Qy 1251 CAGTATTTGACACTTGTGCTTAATGATATCTTCATGACAGACATCCCTGATGAATC 1310
Db 381 SerMetIleProAlaValIleGlnIleIleuArgGlyValAspAspProAlaIleVal 400
Qy 1311 CGAAGCAGCTTTCTGAGCTGTGCTGAGATGTGTTCTTGTGCTCCCTGACATGATCA 1370
Db 401 SerGlnLeuIleLeuAspMetPheGlyAspIlePhePheGlyIleProAlaValLeuLeu 420
Qy 1371 GCTGATATTCAGAGATGCTGTGCTGCTGTCTTCTTATGATTTCCGACCGGCT 1430
Db 421 SerArgSerLeuAlaGlyAspAlaGlyValSerThrIleMetIleGlnPheArgIleArgPro 440
Qy 1431 CAGTCTTTGAGACACAGAGCGGCTTTTGTCAAGCGGACACGCTGATGAATCCGC 1490
Db 441 SerPheValSerAspIleArgProGlnIleThrValGlnIleAspHisGlyAspGlnIlePhe 460
Qy 1491 TTTGTGTTGGGTGGCTCTTCTGAGGGGACATTTGTTTGTGAAAGACCGACGAG 1550
Db 461 PheValPheGlyAlaProLeuLeuLys-----GlnGlyIleSerGln 474
Qy 1551 GAGGAGAAAGTTACTGAGCGGAGAGATGATGAATACTGGGCTACTTTGCTGGAACGGG 1610
Db 475 GlnGlnIleThrAsnLeuSerIleMetValMetIlePheThrPalaAsnPheAlaIleArgAsnGly 494
Qy 1611 AATCTTAATGGAAGCAGCTGTCTCTGTGGCAGCTTATTAATCTGATGAGGATACCTC 1670
Db 495 AsnProAsnGlnGlyGlnIleuProHisIleTrpProGlnIleThrValAspGlnIleGlnIleu 514
Qy 1671 CAGCTGACCTTGACATGAGCTCGGACAGAGACTCAAGAAACCGCGGCTGATTTTGG 1730
Db 515 GlnIleGlyAlaThrThrGlnGlnAlaGlnIleArgLeuValGlnIleValAlaIlePheTrp 534

```

```

Qy 1731 ACCAGACCAATC 1742
Db 535 ThrGlnLeuLeu 538

RESULT 7
S19307
carboxylesterase (EC 3.1.1.1) precursor - pig
N.Alternate names: proline-beta-naphthylamidase
C.Species: Sus scrofa domestica (domestic pig)
C.Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 03-Jul-2004
C.Accession: S19307; S23607
R.Macnushim, M.; Inoue, H.; Ichinose, M.; Tsukada, S.; Miki, K.; Takahashi
FEBS Lett. 293, 37-41, 1991
A.Title: The nucleotide and deduced amino acid sequences of porcine liver proline-beta-n
A.Reference number: S19307; MUD:92070571; PMID:1959668
A.Accession: S19307
A.Molecule type: mRNA
A.Residues: 1-566 <MAT1>
A.Cross-References: UNIPROT:Q29550; EMBL:X63323; NID:g1930; PIDN:CAA44929.1; PID:g1931
A.Accession: S23607
A.Molecule type: protein
A.Residues: 19-40 <MAT2>
A.Note: 28-Lys and 33-Leu were also found
C.Superfamily: cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F.1-18/Domain: signal sequence #status predicted <SIG>
F.19-566/Product: carboxylesterase #status experimental <MAT>
F.51-552/Domain: cholinesterase homology <CHS>
F.80/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.122,467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,72e-80 Length: 566
Score: 1162.50 Matches: 250
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 44.33% Mismatches: 202
Query Match: 30.02% Indels: 41
DB: 2 Gaps: 8

US-10-023-515-1 (1-2158) x S19307 (1-566)
Qy 129 TGTGCTTTTCTGATTTCTCGAGCCCTGTGGACACAGACAGTGGGAAAACTGG 188
Db 2 ThrLeuLeuProLeuValLeuThrSerLeuAlaThrTrp-----AlaGly 19
Qy 189 CTTTGTCTGAAAGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCAAGTC 248
Db 20 GlnProAlaSerProProValValAspThrAlaGlnIleArgValLeuGlyIleVal 39
Qy 249 ACTGTCGCGGAGAGCCCTGTGCTGTGAACGTTCTCGAGAGTCCCTTGTGCTGCTCC 308
Db 40 SerLeuGlnGlyLeuAlaGlnProValAlaValPheLeuGlyValProPheAlaIleAspPro 59
Qy 309 CCGCTGGATCCCTCGATTTACGAACCCGAGAGCTGTGATCGCCCTGGAGATTAATCTGCA 368
Db 60 ProLeuGlnSerLeuAspArgPheAlaIleProGlnIleProAlaGlnIleProTrpSerPheValIle 79
Qy 369 GAAGCACCTCTACCTTAATTGTGCTCCAGAAC-----TCA 407
Db 80 AsnThrThrSerIleProProMetCysCysGlnAspProValValGlnIleMetThrSer 99
Qy 408 CAGTGGCTGCTTATGATCAACACATGCTCAAGTGTCAATCCGAAATTTGCGAGTGTCA 467
Db 100 AspLeuPheThrAsnGlyValGlyIleArgLeuThrLeuGlnIlePhe-----Ser 114
Qy 468 GAAGACTCCCTTACTGATCAATCATATGATGATGATGATGATGATGATGATGATGATG 527
Db 115 GlnAspCysLeuIleValLeuAsnIleThrThrProAlaAspLeuThrIleArgGlyArgLeu 134
Qy 528 CCGGCTTTGGTGTGCTTCCAGAGAGTCCCTTCAAGACTGCTCAGCTCCATCTTGTAT 587
Db 135 ProValMetValTrpIleHisGlyGlyIleValLeuValLeuGlyGlyAlaProMetIleArg 154

```


A:Title: Cloning and sequencing of rat liver carboxylesterase ES-4 (microsomal palmitoyl)
A:Reference number: S62788; MUID:96190723; PMID:8611161
A:Accession: S62788
A:Molecule type: mRNA
A:Residues: 1-561 <ROB>
A:Cross-references: UNIPROT:064573; EMBL:X81825; NID:G550417; PIDN:CAA57419.1; PID:G5504
A:Experimental source: liver
R:Morgan, B.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A:Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID:7986098
A:Accession: S51203
A:Molecule type: protein
A:Residues: 19-48 <MOR>
A:Experimental source: liver
R:Yan, B.; Yang, D.; Brady, M.; Parkinson, A.
J. Biol. Chem. 269, 29688-29696, 1994
A:Title: Rat kidney carboxylesterase. Cloning, sequencing, cellular localization, and re
A:Reference number: A55304; MUID:95050819; PMID:7961958
A:Accession: A55304
A:Molecule type: mRNA
A:Residues: 1-6, 'P', 8-59, 'P', 61-212, 'A', 214-252, 'T', 254-309, 'IT', 312-341, 'N', 343-424, 'P'
A:Cross-references: GB:U10697; NID:G562007
A:Note: the sequence in Genbank entry RNU10697, release 107, (PID:G562008) has the codon
R:Robbi, M.; Beaufay, H.
submitted to the EMBL Data Library, September 1994
A:Reference number: S49257
A:Accession: S49257
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <RO2>
A:Cross-references: EMBL:X81825; NID:G550417; PIDN:CAA57419.1; PID:G550418
C:Function:
A:Description: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-561/Product: carboxylesterase ES-4 #status experimental <Mat>
F:50-551/Domain: cholinesterase homology <CHS>
F:221,466/Active site: Ser, His #status predicted
F:301/Binding site: carboxylate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 1 07e-78 Length: 561
Score: 1139.00 Matches: 243
Percent Similarity: 58.04% Conservative: 82
Best Local Similarity: 43.39% Mismatches: 195
Query Match: 29.42% Indels: 40
DB: 2 Gaps: 10
US-10-023-515-1 (1-2158) x S62788 (1-561)
QY 138 TTCCTGATTCCTCAGCCCTGTGGGACACAGACAGTGGGAAACCTGGCCCTTCTGCT 197
DB 5 PheLeuIleLeuValSerLeuAlaThrCysValValTyrGly-----AsnProSerSer 22
QY 198 GAAAGGCCACAGAGAACACCGAGCTGGATTCAGGGCAAGCAAGTCACTGTGCTG 257
DB 23 Pro---ProValValAspThrThrIlySerGlyLysValLeuGlyLysTyrValSerLeuGlu 41
QY 258 GGAAGCCCTGTCGTGTAAGAGTTCCTCGAGAGTCCCTTGTGCTGCTCCCGCGGGGA 317
DB 42 GlyValIThrGlnSerValAlaValPheLeuGlyValPhePheAlaLysProProLeuGly 61
QY 318 TCCCTGCGATTTACGAACCCGAGAGCTGCATGCGCTTGGGATTAACCTTGAGAGAGCCAC 377
DB 62 SerLeuArgPheAlaLysProGlnProAlaGluProTyrSerPheValLysAsnThrThr 81
QY 378 TCCCTACCTTAATTTGCTCCAGAACTCAAGTGG-----CTGCTCTTA 422
DB 82 ThrTyrProProMetCysSerGlnAspAlaAlaLysGlyGlnArgMetAsnAspLeuLeu 101
QY 423 GATCAACACATGCTCAAGTGCATTAACCGAAATTCGAGAGTGCAGAGACATGCCCTAC 482

DB 102 ThrAsnArgLysGluLysIleLeu-----LeuGluPheSerGluAspCysLeuTyr 118
QY 483 CTGAACATTTATGCGCTGCCACAGCCCATACAGGCTCCCAAGCTCCCGCTTGTGTGG 542
DB 119 LeuAsnIleTyrThrProAlaAspPheThrLysAsnSerArgLeuProValMetValTyr 138
QY 543 TTCACGAGAGTGCCTTCAAGCTGAGCTCAGCTCCATCTTTGATGGATGCGCCGCGGT 602
DB 139 IleHISGlyGlyMetThrLeuGlyGlyAlaSerThrTyrAspGlyArgValLeuSer 158
QY 603 GCGTATGAGAGCTGTGCTGTGTGCTGCTGCTCAATCCGCTAGAAATATTTGTTCTTC 662
DB 159 AlaTyrGluAsnValValValAlaIleGlnTyrArgLeuGlyIleTyrGlyPhePhe 178
QY 663 ACCACATGGAGATCAGATCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGTCTG 722
DB 179 SerThrGlyAspGluHisIleSerArgLysAsnTyrGlyHisLeuAspGluValAlaIleLeu 198
QY 723 TCCGTGGTCCAGAGAACATCGAGTTCTTGGTGGGAGCCCGAGCTGTGACCATCTT 782
DB 199 HisTyrValGlnAspAsnIleAlaAsnPheGlyLysAspProGlySerValIleThrIlePhe 218
QY 783 GCGAGTCCCGCGGAGCCATTAAGTGTTCATCTTATATCTGTCTCCATGGCCAAAGGC 842
DB 219 GlyGluSerAlaGlyGlyPheSerValSerValLeuValLeuSerProLeuThrLysAsn 238
QY 843 TTATTCACAAAGCCATCATCGAGAGTGGGGTGGCCATATCCCTTACCTGAGGCCCAT 902
DB 239 LeuPheHisArgAlaIleSerGluSerGlyValAlaPheLeuProGlyLeuLeuThrLys 258
QY 903 GATTATGAGAGAGAGAGAGAGCTGCGAGTGGTGGACATTTCTGTTGTTAAATGGCTCA 962
DB 259 AspValAlaGlyProAla-----AlaLysGlnIleAlaAspMetAlaGlySerIleThr 276
QY 963 GACTGTAGGCGCTGCTGAGTGGCTTGAAGCAAAACCTCCAGAGAGCTGTGACCCCT 1022
DB 277 ThrSerAlaIleIleValHisCysLeuArgGlnLysThrGluGluGluLeuGluIle 296
QY 1023 ACCGAGAAACAAAG----- 1037
DB 297 MetLysLysMetAsnLeuIleLysLeuSerSerGlnArgAspAsnLysGluSerTyrHis 316
QY 1038 TCTTTCACCTCAGAGTGGTATGATGCTTCTTCTTAATGAGCTCTAGATCATTTGCT 1097
DB 317 PheLeuSerThrValValAspAsnValValLeuProLysAspProLysGluIleAsnAla 336
QY 1098 CAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGAGTCAATACCAAGAGTGGCTTC 1157
DB 337 GluLysAsnPheAsnThrValProTyrIleValGlyIleAsnLysGlnGluCysGlyTyr 356
QY 1158 CTGTGCTCT-----ATGAGAGAGGCTCCTGAGATCTTCATGAGTCCCAAGATCCCTT 1211
DB 357 LeuLeuProThrMetMetC1ylPheValProAlaAspValGlu---LeuAspLysLysMet 375
QY 1212 GCCCTCCATCTGATCAAA-----AACATCTGACATCCCGGCTCGATTAATTGGAC 1262
DB 376 AlaIleThrLeuLeuGlnLysPheAlaSerLeuTyrGlyIleProGluAspIleIlePro 395
QY 1263 CTGTGCTTAATGATATCTTCATGACAAAGCACTCCCTGATCAATCCAGACAGTCTT 1322
DB 396 ValAlaIleGluLysTyrArgLysGlySerAspAspSerIleLysIleArgAspGlyIle 415
QY 1323 CTGACCTGCTTGAAGATGTGTTCTTGTGCTGCTGCTGACATATCACAGCTGATATCAC 1382
DB 416 LeuAlaPheIleGlyAspValSerPheSerIleProSerValMetValSerArgAspHis 435
QY 1383 AGAGATCTGTGAGCACTGCTACTTATGAGTTTGGGACCGGCTCAAGTCTTGA 1442
DB 436 ArgAspAlaGlyAlaProThrTyrMetCtyrGluTyrGlnTyrTyrProSerPheSerSer 455
QY 1443 GACACGAAAGCGGCTTTGTCAAAAGCCGACACAGCTGATGAATCGCTTGTGTGCT 1502

Db 456 ProGlnArgProLyHisValValGlyAspHisAlaAspLeuTyrSerValPheGly 475
QY 1503 GGTGGCTTCCTGAAGGGGACATTTGTATGTTCCAGAGACCGAGAGAGAAAGTTA 1562
Db 476 AlaProIleLeuArg-----AerGlyAlaSerGluGluIleLeu 489
QY 1563 CTGACCCGAGAGATATGAAATACTGGGCTACCTTGTCCGAACCGGGAACTCTAATGCG 1622
Db 490 LeuSerLysMetValMetLysPheThrPalaAsnPheAlaArgAsnGlyAsnProAsnGly 509
QY 1623 AACGACCTGTCTCTGTGGCCAGCTATATATCTGACTGAGACATCTTCAGCTGAGCTTG 1682
Db 510 ArgGlyLeuProHisLTrpProGlnInTyrAspGlnLysGluGluTyrLeuGlnIleGlyAla 529
QY 1683 AACATGACCTCGGACAGACTCAAGAACCGGGGGGTGATTTTGTGACCAAGACCATC 1742
Db 530 ThrThrGlnInSerGlnArgLysLeuValGluGluValAlaPheThrGlnLeuLeu 549
RESULT 11
A55281
cardoxylesterase (EC 3.1.1.1) egaSyn - mouse
N/Alternate names: beta-glucuronidase endoplasmic reticulum-targeting protein; esterase-
C/Species: Mus musculus (house mouse)
C/Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: A55281
R/Ontic, M.; Swanik, R.T.; Fletcher, C.; Zhen, L.; Novak, E.K.; Baumann, H.; Heintz, N.;
Genomics 11, 956-967, 1991
A/Title: Characterization and functional expression of a cDNA encoding egaSyn (esterase-
A/Reference number: A55281, MUID:92147141, PMID:1783403
A/Accession: A55281
A/Status: preliminary
A/Molecule type: mRNA; protein
A/Residues: 1-562 <OVN>
A/Cross-references: UNIPROT:Q64176; GB:880191; NID:9244727; PIDD:AB21335.1; PID:9244728
A/Note: sequence extracted from NCBI backbone (NCBIIN:80191, NCBITP:80194)
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase, endoplasmic reticulum
F/51-552/Domain: cholinesterase homology <CH>
F/222,467/Active site: Ser, His #status predicted
Alignment Scores:
Pred. No.: 1,17e-78 Length: 562
Score: 1138.50 Matches: 240
Percent Similarity: 58.52% Conservative: 76
Best Local Similarity: 44.44% Mismatches: 195
Query Match: 29,408 Indels: 29
DB: 2 Gaps: 6
US-10-023-515-1 (1-2158) x A55281 (1-562)
QY 186 GGGCCCTTCCTGAAGGGGACAGAGAAACCAAGGCTGGATTCAGGGCAAGCA 245
Db 19 GlnHisProSerSerProProMetValAlaPThrValGlnGlyValLeuGlyLysTyr 38
QY 246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCGAGAGCTCCCTTGCTGCT 305
Db 39 IleSerLeuGluGlyPheThrGlnProValAlaValIleLeuGlyValProPheLeuLys 58
QY 306 CCCCCCTGGGATCCCTGCGATTACGAACCCGACGCTGCATGCGCCCTGGATTAACCTG 365
Db 59 ProProLeuGlySerLeuArgPheAlaProProGlnProAlaGluProTyrSerVal 78
QY 366 CGAGAACCCACTCTCAATTTGTCCTCCCAAGCTGAGACTGAGTGGCTCTTATGAT 425
Db 79 LysAsnAlaThrSerTyrProProMetCysPheGlnAspProValThrGlyGlnIleVal 98
QY 426 CAACACATGCTCAAGGTGACTTATCCGAAATTCGAGTG-----TCAGAAAGTGGCTC 479
Db 99 AsnAspLeuLeuThrAsnArgLysGluLysLLeProLeuGlnPheSerGluAspCysLeu 118
QY 480 TACCGAAGCATCTATGGCCCTGCGACGCGCATACAGGCTCCAGCTCCCGTCTTGCTG 539
Db 119 TyrLeuAsnIleTyrThrProAlaAspLeuThrLysSerAspArgLeuProValMetVal 138

QY 540 TGGTCCAGAGAGGTGCTTCAAGACTGGCTCAGGCTCCATGCTTGTGATGGTCCGCTG 599
Db 139 TrpIleHisGlyGlyGlyLeuValLeuGlyGlyAlaSerThrTyrAspGlyLeuValLeu 158
QY 600 GCTGCTATGAGACAGCTGTGGTGTGCTGTGCTGCATGACCGGCTAGAAATATTTGGTTG 659
Db 159 SerThrHisGluAsnValValValValIleGlnTyrArgLeuGlyIleTrpGlyPhe 178
QY 660 TTCACCAATGGGATCAGACTGCTCCGGGAACCTGGGCTTCAGAGACAGAGTGGCTGT 719
Db 179 PheSerThrGlyAspGluHisSerArgLysAsnTrpGlyHisLeuAspGlnValAlaIle 198
QY 720 CTGTCTGTGGTCCAGAGAACATCGAGTTCTTGGTGGAGCCCGACGCTCGTGAACATC 779
Db 199 LeuHisLTrpValGlnAspAsnIleAlaLysPheGlyGlyAspProGlySerValThrIle 218
QY 780 TTTGGCGAGTCCCGGGGACCAATAAAGTGTTCATGCTTATATGCTGTCCATGGCCAA 839
Db 219 PheGlyGluSerAlaGlyGlyGluSerValSerValLeuValLeuSerProLeuAlaLys 238
QY 840 GGCTTATTCACAAAGCCATCATGAGAGTGGGGGCGCATCATCCCTTACAGGCTGAGACC 899
Db 239 AsnLeuPheGlnArgAlaIleSerGluSerGlyAlaLeuThrAlaGlyLeuValLys 258
QY 900 CATGATTATGAGAAAGTGAAGACCTGCAGGCTGTCACATTTCTGTGTAACAATGCG 959
Db 259 LysAsnThrArgProLeuAlaGluLysIleAlaValIleSerGlyCysLysAsn----- 276
QY 960 TCAGACTGTGAGCGCCCTGAGGCTGAGTGAAGAGCAAAACCCCTCAAGAGCTGTGAC 1019
Db 277 ThrThrSerAlaAlaMetValHisCysLeuArgLysThrGluGluGluLeuGly 296
QY 1020 CTCAGCCAGAAACAAAGTCTTT----- 1043
Db 297 ThrThrLeuLysLeuAsnLeuPheLysLeuAspLeuHisGlyAspSerArgGlnSerHis 316
QY 1044 -----ACTGAGTGTGATGATGCTGCTTTCTTCCATGATGAGCTCTAGATCTATTG 1094
Db 317 ProPheValProThrValLeuAspGlyValLeuLeuProLysMetProGluGluIleLeu 1094
QY 1095 TCTCAGAAAGCATTTAAAGCAATCTCTCCATCATGAGTGAATTAACAGAGTGTGGC 1154
Db 337 AlaGluLysAsnPheAsnThrValProTyrIleValGlyIleAsnLysGlnGluPheGly 356
QY 1155 TTCTGTGCTGCT---ATGAAGAGGCTCTTGAGATCTCAGTGGCTCAAGACGCTCTT 1211
Db 357 TrpIleLeuProThrMetMetAsnTyrProProSerAspValLysLeuAspGlnMetThr 376
QY 1212 GCCCTCATGTATACAAAC-----ATCCTGCACATCCCGCTCAGATTTGAC 1262
Db 377 AlaMetSerLeuLeuLysSerSerPheLeuLeuAsnLeuProGluAspAlaIleAla 396
QY 1263 CTGTGGCTTAATGAATATCTTCATGACAGACATCCCTGACAGAAATCCAGACAGTCTT 1322
Db 397 ValAlaIleGluLysTyrLeuArgAspLysAspTyrThrGlyArgAsnLysAspGlnLeu 416
QY 1323 CTGAGCTGTGAGATGATGTTCTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
Db 417 LeuGlnLeuIleGlyAspValAlaPheGlyValIleProSerValIleValSerArgLysHis 436
QY 1383 AAGATGTGTGAGCACTGTCTATGATTTATGAGTTTGGACCGGCTCAGTGGCTTTGAA 1442
Db 437 ArgAspAlaGlyAlaIleProThrTyrMetTyrGluPheGlnTyrSerProSerPheSerSer 456
QY 1443 GACAGAAAGCCGGCTTTGTCAAGCCGACAGCTGATGAAGTCCGCTTGTGTTCGCT 1502
Db 457 GluMetLysProAspThrValValGlyAspHisGlyAspGluIleTyrSerValPheGly 476
QY 1503 GGTGCTTCCTGAAGGGGACATTTGTATGTTCCAGAGGCGACGAGAGAGAAAGTTA 1562
Db 477 AlaProIleLeuArgGly-----GlyThrSerGluGluGluIleAsn 490


```

Oy      1563  CGAGCCGGGAAGATGAATTAAGTCTGAGCTTCTGCAACCGGGAAATCCAAATGG 1622
      |||.....|
Db      491  LeuSerlyseMetMetLeuLysPheTrpAlaasnPheAlaargangLysasnProAsnGly 510
      |||.....|
Oy      1623  AACGACCTGTCTCTGTGGCCAGCTTATATCTGAAGTACAGCACTCCAGCTGACCTTG 1682
      |||.....|
Db      511  GlnGlyLeuProHistrProGlnIuTrpGlnIuLysGlnIuGlyTrpLeuGlnIleGlyAla 530
      |||.....|
Oy      1683  AACATGAGCTTCGGACGACGACTCAAGAACCGGGGTGATTTTGGACGACGACCATC 1742
      |||.....|
Db      531  ThrThrGlnGlnAlaGlnLysLeuLysGlnIuLysGlnIuValAlaPheTrpThrGlnLeu 550
      |||.....|

RESULT 12
A31584
carboxylesterase (EC 3.1.1.1) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-May-1990 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
C:Accession: A31584
R:Long, R.M.; Satoh, H.; Martin, B.M.; Kimura, S.; Gonzalez, F.V.; Pohl, L.R.
Biochem. Biophys. Res. Commun. 156, 866-873, 1988
A:Title: Rat liver carboxylesterase: cDNA cloning, sequencing, and evidence for a multigene
A:Reference number: A31584; MUID:89050119; PMID:2973315
A:Accession: A31584
A:Molecule type: mRNA
A:Releides: 1-540 <LON>
A:Cross-references: UNIPROT:P10959; GB:M20629; GB:X13587; NID:G203279; PIDN:AAA40871.1,
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; cholinesterase; glycoprotein
F:1-9/Domain: signal sequence #status predicted <SIG>
F:10-540/Product: carboxylesterase #status predicted <MAT>
F:41-528/Domain: cholinesterase homology <CHE>
F:70,265,266,293,366,467/Binding site: carboxylate (asn) (covalent) #status predicted
F:212,444/Active site: Ser, His #status predicted

```

[illegible]

Dd	129	TrpIleHisGlyGlyLeuIleIleGlyValAsnProTyrSerGlyLeuAlaIleu	146
Qy	600	GSCTGCAATAGAGACGTGCTGGTGTGGTGCATGCAATCCGGCTAGGAATATTTGGTTCC	659
Dd	149	SerAlaHisGluAsnValValValValThrIleGlnTyrThrGluGlyPheGlyGlyLeu	168
Qy	660	TTTCAACCAATGGGATTCAGACATGCTCCGGGGAACTGGGCTTTCAAGAACAGGTGGCTCT	719
Dd	169	PheSerThrGlyAspGlnHisSerAspGlyAsnTrpAlaHisLeuAspGlnLeuAlaAla	188
Qy	720	CTGTCCTGGGCTCCAGAAAGACATGAGCTCTTCGGTGGGGACCCACACTCTGTACCATC	779
Dd	189	LeuArgTrpValGlnAspAsnIleAlaAsnPheGlyValAsnProAspSerValThrIle	208
Qy	780	TTTGGCGAGTCCGGGAGCCATTAAGTGTTCATGCTTATAGTCTCCATGGCCAAA	839
Dd	209	PheGlyGlnSerAlaGlyGlyValSerValSerAlaLeuValLeuSerProLeuAlaIle	228
Qy	840	GGCTTATTTCCAAAGCCATCATGAGAGTGGGGTGGCCATCATCTCTTACCTTGAGGCC	899
Dd	229	AsnLeuPheHisArgAlaIleSerGlnSerGlyValValIleThrHisAsnLeuAspIle	248
Qy	900	CATGATTTATGAGAGATGAGAGCACTCGACGGTGGTTCACATTTCTGTGGTAACAATGCG	959
Dd	249	LysAsnThrGlnAlaVal-----AlaGlnMetIleAlaThrLeuSerGlyCysAsnAsn	266
Qy	960	TCAAGCTCTGAGCGCCCTGAGGTGGCTGAGAGCAAAACCTTCCAAAGCTGCTGACC	1019
Dd	267	ThrSerSerAlaAlaMetValGlnCysLeuArgGlnIlyThrGlnAlaGlnLeuLeuGln	286
Qy	1020	CTCAGCCAGAAA-----ACAAAGCTTTTCACTCGAGTGGTGAATGGTCTTTCTTCT	1073
Dd	287	LeuThrValValValLeuAspAsnThrSerMetSerThrValIleAspGlyValIleLeuPro	306
Qy	1074	AATAGGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCCATCATGGA	1133
Dd	307	LysThrProGlnGlnIleLeuThrGlnIlySerPheAsnThrValProTyrIleValGly	326
Qy	1134	GTCATATAACCAAGAGTGGCTGCTCTGCTGCTGATGAAGAGGCTCTGATGATCTCAGT	1193
Dd	327	PheAsnIlySerGlnIlyPheGlyTrpIleIleProThrMetCysIleAsnLeuLeuSerGln	346
Qy	1194	GGC---TCCAAAGATGCCCTTGGCCCTCATCTGATCAAAAACATC-----CTGCAC	1241
Dd	347	GlyThrMetAsnGlnIlyMetAlaIleSerSerPheLeuValArgPheSerProAsnLeuAsn	366
Qy	1242	ATCCCGGCTCAGATTGTGGACCTTGGCGTAAATGATCTTCCATGACAAAGACTCCCTG	1301
Dd	367	IleSerGlnSerValIleProAlaIleIleGlnIlyThrIleuArgIlyThrAspAspPro	386
Qy	1302	ACTGAATTCGAGACAGTCTTGGGACCTTGGCATGCGTTCGATAGTTCCTTGGTCCCTGCA	1361
Dd	387	AlaIlyValValValValLeuLeuLeuAspMetPheSerAspValPhePheGlyIleProAla	406
Qy	1362	CTGATTCACAGTCCATATTCACAGAGAGTGGTGGTGCACCTGTACTTATGATTTCCG	1421
Dd	407	ValLeuMetSerAlaGSerIleuAspAlaGlyAlaIleProThrIlyMetCylArgIlyPheGln	426
Qy	1422	CACCGGCTCAGTGCCTTTGAAGACAGAGCCGGCTTTTTCAAAGCCACACGCTGAT	1481
Dd	427	TyrArgProSerPheValSerAspGlnArgProGlnThrValGlnIlyAspHisGlyAsp	446
Qy	1482	GAAATCCGCTTGGTTCGGTGGTGGCTCTGTAAGGGGAGCATTTGTTATGTTCCAAAG	1541
Dd	447	GlnIlePheSerValPheGlyThrProPheLeuIys-----GlnGly	460
Qy	1542	GCCACGAGAGAGAGATTAAGTACGGCCGAGAGATGTGAATTAAGTGGCTACCTTGGCT	1601
Dd	461	AlaSerGlnGlnIlyThrIleAsnLeuSerIlyLeuValIleMetIlySerPheTrpAlaAsnPheAla	480
Qy	1602	CGAAACGGGAATCTTAATGGGAACGACTGTCTCTGTGGCCAGCTTAATCTGACTGAG	1661

OY		1074	AATAGACCTCAGATCTATTGTCTCCAGAAAGCATTTTAAGAACAATTCCTTCATCATCGGA	1133
			: : : : : : : : : : : : : : : : : : : :	
Dd		307	LysThrProGluGlnIleLeuThrGlnLysSerPheHisnThrValProTyrIleValGly	326
OY		1134	GTCATAAACCAGAGTGTCGCTTCCTGCCTGACTATGAAGAGGCTCCTGAGATCTCACT	1193
			: : : : : : : : : : : : : : : : : : : :	
Dd		327	PheAsnLysGlnGlnPheGlyTyrPileIleProThrMetMetGlyAsnLeuSerGln	346
OY		1194	GGC---TCCAACAAGTCCCCTGCCCTCATCTGATACAAAACATC-----CTGCAC	1241
			: : : : : : : : : : : : : : : : : : : :	
Dd		347	GlyArgMetAsnGlnLysMetAlaSerSerPheLeuLysArgPheSerProAsnLeuAsn	366
OY		1242	ATCCCGGCTCAGTATTGGACACTTGTCGCTAATGAATACCTTCATGACAGACATCCCTG	1301
			: : : : : : : : : : : : : : : : : : : :	
Dd		367	IleSerGlnSerValIleProAlaIleIleGlnLysTyrLeuArgGlyThrAspAspPro	386
OY		1302	ACTGAATAATCCAGACACATCTCTTCGAGACTTCCTGAGAGATGTTCTTTGTCGCCGCA	1361
			::: : : : : : : : : : : : : : : : : : : :	
Dd		387	AlaLysIleLysIleGlnLeuLeuLeuAspMetPheSerAspValPhePheGlyIleProAla	406
OY		1362	CTGATCAACAGCTGATATATCACAGAGATGTCGTGCGACCTGCTCTATCTTATGACTTCCG	1421
			::: : : : : : : : : : : : : : : : : : : :	
Dd		407	ValLeuMetSerArgSerLeuArgAspAlaGlyAlaIleProThrTyrMetTyrGlnPheGln	426
OY		1422	CACCGGCTCAGTCGCTTTGAAGACACGAACCGCGCTTTTGCAAAGCCGACCAACGCTGAT	1481
			::: : : : : : : : : : : : : : : : : : : :	
Dd		427	TyArGrProSerPheValIserAspGlnArgProGlnThrValGlnGlyAspPheIleGlyAsp	446
OY		1482	GAAATCCGCTTTGTTGTCGATGCGATGCGCTTCCTCGAAGGGGACATGTTGTTATGTTGAAAGA	1541
			: : : : : : : : : : : : : : : : : : : :	
Dd		447	GlnIlePheSerValPheGlyThrProPheLeuLys-----GlnGly	460
OY		1542	GCCACGAGAGAGAGAATTACTGAGCCGCGAAGATGATGAATACTGGGCTACCTTGTCT	1601
			: : : : : : : : : : :: : : : : : : : : : :	
Dd		461	AlaSerGlnGlnGlnThrAsnLeuSerLysValMetLysPheTrpIleAsnPheAla	480
OY		1602	CGAACCGGGAATCTTAATGGGAACGACCTGTCTCTGCGCAGCTTAAATCTGACTGAG	1661
			: : : : : : : : : : : : : : : : : : : :	

D _b	461	ArgaenGlyAsehpProAsnGlyClnGlyLeuProH1StrProlYstyRspGlnIysGlu	500
O _y	1662	CAGTACCTCCAGCTGGACTGAACAATGAGCCCTGGACAGAGACTAAAGAACC GGCGGTG	1721
D _b	501	GlyTyrlleuGlnIIeGIyAlaThrThrgInGlnIaGlnIysleuIysGlyGlnGluVal	520
O _y	1722	GATTTTTGACACGACGACCATC	1742
D _b	521	AlaphetTrfhnGlnIuleu	527

QY 1611 AATCTAATGGGAACGACCTGTCTCTGTCGACGTTATATCTGACTGACGATACCTC 1670
|||||
Db 506 AaenProbaenAlaArgGlyLeuProHisTrpProGlnTrpAspGlnLysGlnLysTrpLeu 525
|||||
QY 1671 CAGCTGACCTTGACACATGAGCTCTCGACAGACATCAAGAAACCGCGGTGATTTTGG 1730
|||||
Db 526 GlnIleGlyAlaThrThrGlnInSerGlnArgLeuValGlnLysAlaIlePheTrp 545
|||||
QY 1731 ACCGACCAATC 1742
|||
Db 546 ThrGlnLeuLeu 549
|||
RESULT 14
JC2447
carboxylesterase (EC 3.1.1.1) ES-3 precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 29-Mar-1995 #sequence _revision 26-May-1995 #text_change 09-Jul-2004
C/Accession: JC2447; S23462
R/Robbi, M.; Beaufay, H.
Biochem. Biophys. Res. Commun. 203, 1404-1411, 1994
A/Title: Cloning and sequencing of rat liver carboxylesterase ES-3 (EGASTN).
A/Reference number: JC2447; MUID:95032008; PMID:7945287
A/Accession: JC2447
A/Molecule type: mRNA
A/Residues: 1-561 <ROB>
A/Cross-references: UNIPROT:063108; GB:X81395; NID:9550146; PIDN:CAA57158.1; PID:9550147
A/Experimental source: Liver
R/Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A/Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A/Reference number: S23460; MUID:92299008; PMID:1606962
A/Accession: S23462
A/Status: preliminary; translation not shown
A/Molecule type: mRNA
A/Residues: 503-554, 'R', 556-561 <MED>
A/Cross-references: EMBL:X65295; NID:957557; PIDN:CAA46390.1; PID:957558
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; glycoprotein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-561/Product: carboxylesterase ES-3 #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CHB>
F:79,107,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:221,466/Active site: Ser, His #status predicted
Alignment Scores:
Pred. No.: 1.78e-77 Length: 561
Score: 1123.00 Matches: 245
Percent Similarity: 57.07% Conservative: 86
Best Local Similarity: 42.24% Mismatches: 193
Query Match: 29.00% Indels: 56
DB: 2 Gaps: 11
US-10-023-515-1 (1-2158) x JC2447 (1-561)
QY 132 TGCTTTTCCTGATTCCTCAGCCCTGTGGACACACAGTGGGAAAACTGGCCT 191
|||
Db 2 CysLeuTrpAlaLeuIleLeuValPheLeuAlaIlePheThrAlaGly-----GlyHis 19
|||
QY 192 TCTGTGAAGGCGCCACAGACAGACAGGCTGGATGATTCAGGACGACATCACT 251
|||
Db 20 ProSerSerLeuProValAlaPheThrLeuGlnGlyLysValLeuGlyLysTrpValSer 39
|||
QY 252 GTGCTGGAGACCTGTGCTGTGTAACGTGTTCTCGAGATCCCTTGTGCTGCTCCCG 311
|||
Db 40 LeuGlnGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaLysProPro 59
|||
QY 312 CTGGATCCCTCGGATTTACGAACCCGACGCTGATGCTCCCTGGGATTAATCTGGAGAA 371
|||
Db 60 LeuGlySerLeuArgPheAlaProProGlnProAlaGlnProTrpSerPheValLysAsn 79
|||
QY 372 GCCACCTCTCAATTTGTCCTCCAGAAC----- 404
|||

Db 80 ThrTrpSerTrpProPheMetCysSerGlnAspProValAlaGlyGlnIleValAsnAsp 99
QY 405 -----TCAGAGTGGCTGCTTATGATCAACAGCATGCTCAAGGTGATTAACCGAAATTC 458
|||
Db 100 LeuLeuTrpAsnTrp-----GlnGlnAsnIleSerLeuGlnPhe----- 112
|||
QY 459 GGAGTGTCAAGACACTGCTCTACCTGAACATCTATGCGCTCCGACCGCATACAGGC 518
|||
Db 113 -----SerGlnAspCysLeuTrpLeuAsnIleTrpThrProAlaAspLeuThrLysArg 130
|||
QY 519 TCAGAGCTCCCGCTTGTGTTGTTCCAGAGAGTCCCTTAAGATCGCTCAGCTTC 578
|||
Db 131 AspArgLeuProValIleMetValTrpIleHisGlyGlyGlyLeuValLeuGlyLysAsp 150
|||
QY 579 ATCTTTGATGGGTCCGCTGCTGCTCCATATGAGAGAGCTGTTGTGGTCCGATGCT 638
|||
Db 151 ThrTrpAspGlyLeuAlaLeuSerThrHisGlnAsnValValValValIleGlnTrp 170
|||
QY 639 CGGCTAGGAAATATTTGGTTTCTTACACACATGGAGTCAAGATGCTCCGCGGAACTGGCC 698
|||
Db 171 ArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerArgGlyAsnTrpGly 190
|||
QY 699 TTCAAGACCAAGTGGCTGCTGTCTGCTGCTCCAGAAACATGAGTCTTCGATGG 758
|||
Db 191 HisLeuAspGlnValAlaAlaLeuHisTrpValGlnAspAsnIleAspAsnHeGlyGly 210
|||
QY 759 GACCCCACTCTGTACATCTTTGGCGAGTCCGCGGAGCATTAAGTTCTTATGCTCT 818
|||
Db 211 AspProGlySerValThrIlePheGlyGlnSerAlaGlyGlyGlnSerValSerValLeu 230
|||
QY 819 ATATGCTCTCCATGCGCAAGAGCTTATTCACAAAGCATCATGAGATGGAGTGGCC 878
|||
Db 231 ValLeuSerProLeuAlaLysAsnLeuPheHisLysAlaIleSerGlnSerGlyValAla 250
|||
QY 879 ATCATCCCTTAAGTCTG--GAGGCCATGATTAAGAAAGTGAAGACCTGCAAGTGGT 935
|||
Db 251 LeuThrAlaGlyLeuValLysLeuAsnThrArgProLeuAlaGlnLysIleAlaVal 270
|||
QY 936 GCACATTTCTGTGTACAAATCCCTCAAGCTCTGAGGCCCTGAGGCTCGTGGAGAA 995
|||
Db 271 Ser-----GlyCysLysSerThrThrSerAlaSerMetValHisCysLeuArgGln 287
|||
QY 996 AAACCTCAAGAGACTGCTGACCTTCCAGAAACAAAGTCTTTCATC----- 1046
|||
Db 288 LysThrGlnGlnGlnLeuLeuGlnThrThrLeuLysLeuAsnLeuPheSerLeuAspLeu 307
|||
QY 1047 -----CGAGTGTGATGTCCTTTCTT 1070
|||
Db 308 HisGlyAspSerArgGlnSerTrpProPheValProThrValLeuAspGlyValValLeu 327
|||
QY 1071 CCTAATGAGCTCTAGATCTATGTTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATC 1130
|||
Db 328 ProLysMetProGlnGlnIleLeuAlaGlnLysAspPheAsnThrValProTrpLysVal 347
|||
QY 1131 GGAGTCAATTAACAGAGTGTGCTTCCGCTGCT-----ATGAAGAGAGCTCTGAGATC 1187
|||
Db 348 GlyIleAsnLysGlnGlnPheGlyTrpIleLeuProThrMetCysAsnTrpProProSer 367
|||
QY 1188 CTCAGTGGCTCCACAGATGCTCCCTGCTCCATCTGATACAAAC-----ATCTG 1238
|||
Db 368 AspMetLysLeuAspProMetThrAlaThrSerLeuLysLysSerSerPheLeuLeu 387
|||
QY 1239 CACATCCGCTCAGATATTTGACCTTGTGGTAAATGAATCTTCATGACAGACATCC 1298
|||
Db 388 AsnLeuProGlnGlnAlaIleProValAlaValGlnLysTrpLeuArgHisThrAspAsp 407
|||
QY 1299 CTGACTGAATCCAGACAGTCTTGTGAGCTTGTGAGATGTGTTGTGCTGCTCT 1358
|||
Db 408 ProAspArgAsnLysAspGlnLeuLeuGlnIleGlyAspValIlePheGlyValPro 427
|||
QY 1359 GCACGTATCAGACGCTGATATCAGACAGATGCTGGTGCACCTGTCTATGAGTTT 1418
|||
Db 428 SerValIleValSerArgGlyHisArgAspAlaGlyAlaGlyThrTyMetTrpGlnPhe 447
|||

QY 1419 CGGACACGGCCCTCAGTGGCTTTGAAAGACAGAACCGCGGCTTTTTCGAAAGCCGACAGCCT 1478
 Db
 448 GlnTyrArgProSerPheSerSerIleMetLysProSerThrValValGlyAspHisGly 467
 QY 1479 GATGAAGTCCGCTTGTGTGGTGGTGGTCCCTTCCGAAAGGAGGACATTGTATTGTTGAA 1538
 Db
 468 AspGlnIleTyrSerValPheGlyValAlaProIleLeuThrGly----- 481
 QY 1539 GAGCCACGAGAGAGAGAAAGTACTGAGCCGGAAGATGATGAATAATCGGCTTACT 1598
 Db
 482 GlyThrSerLysGluGlnIleAsnLeuSerIleMetMetLysPheTrpAlaAsnPhe 501
 QY 1599 GCTCCGACCCGGGAATCTTAATGGAGACAGCTGTCTCTGTGGCCAGCTTTAATTCGACT 1658
 Db
 502 AlaArgAsnGlyAsnProAsnGlyGlnIleuProHisTrpProGlnTyrAspGlnLys 521
 QY 1659 GAGCAGTACCTTCAGCTGAGTGAATGAACATGAGCCTCGACAGAGACTCAAGAGACCGCG 1718
 Db
 522 GlnGlyTyrLeuGlnIleGlyAlaThrThrGlnGlnAlaGlnIleuLysGlnLysGlu 541
 QY 1719 GTGATTTTGGACACGACCATCCCTGATCTGTCTGTCCGACATGCTCCAGCT 1778
 Db
 542 ValAlaIleThrPseGlu-----LeuLeuAlaMetLysProLeuHisAla 556

RESULT 15
 S10367
 carboxylesterase (EC 3.1.1.1) ES-10 precursor, microsomal - rat
 N/Alternate names: hydrolase A
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #ext_change 09-Jul-2004
 C/Accession: S10367; S12468; S51202; S23460; S14361
 R/Robb: M.; Beaufay, H.; Octave, J.N.
 Biochem. J. 259, 451-458, 1990
 A>Title: Nucleotide sequence of cDNA coding for rat liver pi 6.1 esterase (ES-10), a carboxylesterase
 A/Reference number: S10367; MUID:90351366; PMID:2386485
 A/Accession: S10367
 A/Molecule type: mRNA
 A/Residues: 1-565 <ROB1>
 A/Cross-references: UNIPROT:Q9R135; EMBL:X51974
 A/Note: 168-Gln, 247-Lys, 423-Met, and 506-Asn were also found
 A/Note: the sequence from P19. 4 is inconsistent with that from Fig. 5 in having 265-Lys
 R/Robb: M.
 submitted to the EMBL Data Library, February 1990
 A/Reference number: S12468
 A/Accession: S12468
 A/Molecule type: mRNA
 A/Residues: 1-264 'K', 266-565 <ROB2>
 A/Cross-references: EMBL:X51974; NID:G56898; PIDN:CAA36236.1; PID:G56899
 R/Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
 Arch. Biochem. Biophys. 315, 495-512, 1994
 A>Title: Purification and characterization of two rat liver microsomal carboxylesterases
 A/Reference number: S51202; MUID:95077430; PMID:7986098
 A/Accession: S51202
 A/Molecule type: protein
 A/Residues: 19-48 <MOR>
 R/Medda, S.; Proia, R.L.
 Eur. J. Biochem. 206, 801-806, 1992
 A>Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
 A/Reference number: S23460; MUID:92299008; PMID:1606962
 A/Accession: S23460
 A/Status: preliminary; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-185 'W', 187-422 'W', 424-505 'N', 507-565 <MED>
 A/Cross-references: EMBL:X65296; NID:G57553; PIDN:CAA46391.1; PID:G57554
 Biochem. J. 274, 693-697, 1991
 A>Title: Purification and characterization of carboxylesterases from rat lung.
 A/Reference number: S14361; MUID:91190080; PMID:2012599
 A/Accession: S14361
 A/Molecule type: protein
 A/Residues: 19-26 'D', 28-37 <GAD>
 C/Superfamily: cholinesterase, cholinesterase homology

C-Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F.1-18/Domain: signal sequence #status predicted <Sig>
F.19-565/Product: carboxylesterase #status predicted <MAM>
F.50-551/Domain: cholinesterase homology <CH>
F.79,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.221,466/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	2,32e-77	Length:	565
Score:	1121.50	Matches:	245
Percent Similarity:	56.06%	Conservative:	79
Best Local Similarity:	42.39%	Mismatches:	199
Query Match:	28.96%	Indels:	55
DB:	2	Gaps:	11

US-10-023-515-1 (1-2158) x S10367 (1-565)

```
QY      129  TGGTGCCTTTTCCTGATTCCTCCAGCCCGCTTGGGACACAGACAGTGGGAAAAATCGG 188
Db      8  TriplePhe-----LeuAlaAlaCysThrAlaTrpGlyTyr----- 19
QY      189  CTTCTGCTGAAGGCCACAGAGAACACAGCTGGATGATTCAGGGACAGCAAGTC 248
Db      20  ProSerSerPro---ProValValanthrVallysglylysevalLeuglylyertyVal 38
QY      249  ACTGAGCTGGGAAGCCCTGTGCTGTGAAGTGTTCCTGGAGTCCCTTTGCTGCTCC 308
Db      39  AenleuGlyGlyPheAlaGlnProValAlaValPheLeuGlyIleProPheAlaIyPro 58
QY      309  CGCGTGGAGTCCCTCGATTTACGAACCCGAGAGCTGTGATCGCCCTGGAGAACTTCGA 368
Db      59  ProleuGlySerleuAlaArgPheAlaIaProGlnProAlaGlnProItrPheAlaPheVal 78
QY      369  GAAGCAGCTCTTCACTCAATTGTCCTCCAGAAC-----TCA 407
Db      79  AenThrThrSerTyrProProMetCysSerGlnAenAlaValIgllyGlnValLeuSer 98
QY      408  GAGTGGCTGCTCTTGAATGATCAGACATGCTGCAAGTGCATTAACCGAAATTCGGAGTCTCA 467
Db      99  GltleuPheThrAenAlaGlylysglnAenIleProleuGlnPhe-----Ser 113
QY      468  GAAGCTGCTCTTCACTGGAACATGATGACGCTGCCACGCCGATACAGGCTCCAGAGTCTC 527
Db      114  GlnAerCysleuTyrLeuAenValTyrThrProAlaSerleuThrIyAaenSerAlaGly 133
QY      528  CCGCTGTGGTGTGGTCCACAGAGAGTCCCTTCAAGACTGCGCTCAGCTTCATTTGAT 587
Db      134  ProValMetValTyrIleNisgIyGlyIyLeuValIgllyIyAlaSerThrTyrAer 153
QY      588  GGGTCCGCCCTGGCTGCTATGAGAGAGTGTGGTGTGTGCTGCCAGTAACGGCTCAAGGA 647
Db      154  GltynIalValLeuSerAlaAlaHsgIuAenValIyValValThrIleGlnTyrArgLeuGly 173
QY      648  ATATTGGATTCTTCAACCAATGGAATCAGACATGCTCCGGAGAACTGGGCTTTCAAGAG 707
Db      174  IleThrIyPhePheSerThrGlyAerGlnHsaSerAlaGlyAenThrIyGlyIyAenAer 193
QY      708  CAGGTGCTGCTGTGCTGCTGGGTCAGAAAGAACATGAAGTCTTTCGGTGGGACCCCAAC 767
Db      194  GlnValAlaAlaLeuNisIyTrpValIglAaPheAlaIleAlaAaPheNisgIyIyAaPheGly 213
QY      768  TCTGTGACCATCTTTGGCGAGTCCGCGGAGCCATTAAGTGTTCATGTTATACGTCT 827
Db      214  SerValThrThrIlePheGlyIuSerAlaIgllyGlyPheSerValSerAlaLeuValLeuSer 233
QY      828  CCATGCGCAAAGGCTTATTTCCAAAGCCATCATGAGAGTGGGGTGGCCATTCCTCT 887
Db      234  ProleuAlaIyAaenleuPheNisIyArgAlaIleSerGlnSerGlyIyValIleuThrSer 253
QY      888  TACCTGAAGGCCCATATTATTAAGAAAGTAGAGACCTGCGAGTGTGGTGCATTTCTGT 947
Db      254  AlaIeuIleThrThrAaPserIySerProIleAlaAaenleu-----IleAlaThrLeuSer 271
```


This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

this Page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 02:39:48 ; Search time 192.639 Seconds

(without alignments)
11472.925 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 3872

Sequence: 1 ccaacgcctccgaataaacag.....aaaaaaaaaaaaaaaaa 2158

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p,model-DEV=xlh
-Q=/cgn2.1/USPTO.epool/6664091/runal_14062005.133459.9335/app.query.fasta_1.4238
-DB=UniProt_03 -QFMT=fasta -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCGALIGN=200 -THR SCORE=dpct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091.cgn.1.1.389@runal_14062005.133459.9335 -NCPU=6 -ICPU=3
-NO MMAP -LARGESOURCE -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03:~*
1: UniProt_sprot:~*
2: UniProt_trembl:~*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2915	75.3	525	2	Q6NT32
2	2606	67.3	575	2	Q96DN9
3	2454	63.4	469	2	Q8N8C8
4	2250	58.1	575	2	Q6AM47
5	2094.5	54.1	575	2	Q6AM46
6	1985	51.3	545	2	Q81034
7	1691.5	43.7	361	2	Q95KH3
8	1298.5	33.5	557	1	SASB ANAPL
9	1282.5	33.1	561	2	Q91WGO
10	1281	33.1	559	2	Q8BK48
11	1276.5	33.0	559	2	Q35533
12	1271	32.8	559	2	Q8BM97
13	1266.5	32.7	556	2	Q6PD87
14	1263.5	32.6	568	2	Q6GM54
15	1262.5	32.6	561	2	Q70177
16	1260.5	32.6	560	2	Q8R097

17	1257.5	32.5	554	2	Q35535
18	1239.5	32.0	562	2	Q6P306
19	1238	32.0	558	2	Q8K3R0
20	1232.5	31.8	558	2	Q8QZ83
21	1231.5	31.8	532	1	EST2_RABIT
22	1230	31.8	559	1	EST2_HUMAN
23	1213.5	31.3	561	2	Q70631
24	1213.5	31.3	566	2	Q766D7
25	1213.5	31.3	566	2	Q64459
26	1210.5	31.3	557	2	Q64076
27	1199	31.0	561	1	EST1_MESAU
28	1191	30.8	586	2	Q61PK9
29	1187.5	30.7	568	2	Q6RG74
30	1187.5	30.7	572	2	Q810S9
31	1186.5	30.6	554	1	ESTM_MOUSE
32	1183	30.6	565	2	Q95N05
33	1179.5	30.5	571	2	Q6UW8
34	1179	30.4	566	2	Q726J1
35	1175	30.3	565	2	Q8RD29
36	1174.5	30.3	568	2	Q8VCU1
37	1169.5	30.2	561	2	Q91WU0
38	1164	30.1	565	2	Q97582
39	1162.5	30.0	555	2	Q6GMJ1
40	1162.5	30.0	566	1	EST1_PIG
41	1159.5	29.9	566	2	Q46421
42	1150.5	29.7	565	1	EST1_RABIT
43	1150	29.7	554	1	ESTN_MOUSE
44	1149	29.7	565	2	Q35534
45	1144	29.5	549	1	EST1_RAT

ALIGNMENTS

RESULT 1	Q6NT32	PRELIMINARY;	PRT;	575 AA.
ID	Q6NT32;			
AC	Q6NT32;			
DT	05-JUL-2004 (Tremblrel. 27, Created)			
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)			
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)			
DE	FLJ31547 protein.			
GN	Name=FLJ31547;			
OS	Homo sapiens (Human) ;			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PCR rescued clones;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,			
RA	Alteich S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Martins K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,			
RA	Brownstein M.J., Ueda T.B., Toshiyuki S., Carrinci P., Prange C.,			
RA	Rah S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Boeser S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villaalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PCR rescued clones;			

RA Director MGC Project;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; BC069501; AA069501.1; -.
DR HSSP; P12337; IK41.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR00379; Ser. ester.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 575 AA; 63926 MW; 6F5B735BDFC9C09 CRC64;

Alignment Scores:

Pred. No.:	4,51e-206	Length:	575
Score:	2915.00	Matches:	553
Percent Similarity:	97.54%	Conservative:	3
Best Local Similarity:	97.02%	Mismatches:	6
Query Match:	75.28%	Indels:	8
DB:	2	Gaps:	1

US-10-023-515-1 (1-2158) x Q6NT32 (1-575)

```
QY 129 TGGTGTCTTTTCCGATTCACAGCCCTGTGGGACACAGACGTGGGAAAACTGGG 188
DB 14 TTPAlaIleTPrValLeuAlaAlaPro-----ThrlYsGly 25
QY 189 CTTTGTGCTGAAGGGCCACAGAGAACACAGGCTGGAGTGAATGAGGCGAAGCAATG 248
DB 26 ProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTTrpIleGlnGlyVal 45
QY 249 ACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTCTCGAGTCCCTTGTGCTGCC 308
DB 46 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaPro 65
QY 309 CCGCTGGGATCCCTGGGATTTAGAAACCGGACGCTGCATCCCGCTGGGATTAATTCGA 368
DB 66 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTrpAsnLeuArg 85
QY 369 GAAGCCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTATGATCA 428
DB 86 GluAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGluTrpLeuLeuAspGln 105
QY 429 CACATGCTGAAGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTCTTACTGTAAC 488
DB 106 HisMetLeuValValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeuAsn 125
QY 489 ATCTATGCGCTGCCACCGCGATACAGGCTCAGAGTCCCGCTTGTGTGCTGCCA 548
DB 126 IleTyrAlaProAlaHisIleAlaPheThrGlySerTyrSLeuProValLeuValTrpPhePro 145
QY 549 GGAGGTGCTTCAAACTGCTCAGCTTCACCTTTGATGGAGTCCGCTGGCTGCTCAT 608
DB 146 GlyGlyAlaPheLeuTyrThrGlySerIleAserIlePheAspGlySerAlaLeuAlaIleTyr 165
QY 609 GAGGACGTGCTGTGTGTGCTGTCAGATACCGGCTAAGGAATTTGGTTTTTACACACA 668
DB 166 GluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 185
QY 669 TGGGATCGCATGCTCCGCGGAACGTGGCTTCAAGACACAGTGGCTGCTGCTGG 728
DB 186 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrp 205
QY 729 GTCCAGAAAGACATCGAGTTCTTGGGTGGGAGACCCGACGCTTGTGACATCTTTGGCAG 788
DB 206 ValGlnTyrAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 225
QY 789 TCCGCGGAGGCATTAAGTGTTCATGCTTAACTGCTCCCATGGCCAAAGGCTTATTC 848
DB 226 SerIleGlyAlaIleSerValIleSerLeuIleLeuSerProMetAlaIleGlyLeuPhe 245
QY 849 CACAAAGCCATCATGAGAGATGGGGTGGCCATATCCCTTACTGAGGCCCATGATTAAT 908
```

```
DB 246 HisIleAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyr 265
QY 909 GAGAAAGTGAAGACCTCGAGGCTGGTGACATTTCTGTGTACATAGCTCAGACTCT 968
DB 266 GluLysSerGluAspLeuGlnValAlaHisPheCysGlyAsnAlaSerLeuSer 285
QY 969 GAGGCCCTGCTGAGGTGCTGAGCAAAACCCCTCAAGAGCTGTACCTCAGCCAG 1028
DB 286 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGln 305
QY 1029 AAAACAAAGCTCTTCACTCGAGTGGTATGATGCTTTCTTCTTAATGAGCTCTGAT 1088
DB 306 LysThrLysSerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProLeuAsp 325
QY 1089 CTATGTCTCGAAGACATTTTAAAGCAATTCCTTCCATCATCGAGTCAATACCCAGAG 1148
DB 326 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHisGlu 345
QY 1149 TGTGCTTCTGCTGCTGCTATGAGAGGCTCTGAGATCCTCAGTGGCTCCAAAGTCC 1208
DB 346 CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLysSer 365
QY 1209 CTTGCGCTCCATCGATTAACAAACATCTCGACATCCCGCTCAGATATTGACCTTTGTG 1268
DB 366 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisLeuVal 385
QY 1269 GCTAATGATATTCATGACACAAAGCACTCCCTGACTGAATTCGAGACAGTCTTGGAC 1328
DB 386 AlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAsp 405
QY 1329 TTTGCTTGAAGATGTGTTCTTTGTGCTGCTGCTGACACTGATACAGCTCGATATCAAGAGAT 1388
DB 406 LeuLeuGlyAspValPhePheValAlaProAlaLeuIleThrAlaArgTyrHisLysAsp 425
QY 1389 GCTGTGTCACCTGTCTTCTTATGAGTTTCGACACCGGCTCAGAGCTTGAAGACAG 1448
DB 426 AlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAspThr 445
QY 1449 AAGCGGCTTTTGTCAAAGCCGACACGCTGATGAAGTCCGCTTGTGCTGGTGGTGGCC 1508
DB 446 LysProAlaPheValIleValAlaAspHisAlaAspGluValArgPheValPheGlyGlyAla 465
QY 1509 TTCCTGAAGGGAGGACATTTGTTATGTTGGAAGGACCGACGAGAGAGAACTTACTGAC 1568
DB 466 PheLeuLysGlyAspIleValIleMetPheGluGlyAlaThrGluGlnGluLysLeuLeuSer 485
QY 1569 CGGAAGATGATGAATATCTGGGCTACCTTGTGCTCGAACCGGGAATCTTAATGGAAACGAC 1628
DB 486 ArgLysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 505
QY 1629 CTGTCTGTGGCGACGCTTAATAATCTGACTGACGCTACCTTCAGCTGACTTGAATG 1688
DB 506 LeuSerLeuTrpProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMet 525
QY 1689 AGCTTCGACAGACACTCAAGAACCGCGGGTGAATTTTGGACGACAGACATCCCGCTG 1748
DB 526 SerLeuGlyGlnArgLeuLysGluProArgValAspPheTrpHisSerThrIleProLeu 545
QY 1749 ATCCGTGTCTGCTCCGACATGCTCCACAGTCCCTTTTCTTAATCTTCTCTCTC 1808
DB 546 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 565
QY 1809 CTCGAGCTTTCTTTTCTTTTGTGCTGCT 1838
DB 566 LeuGlnProPhePhePhePheCysAlaPro 575
RESULT 2
Q66DN9 PRELIMINARY; PRT; 525 AA.
AC Q66DN9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
```



```

Db      |||||
266  GIUYSerGluAspLeuGlnValAlaHisPheCySGlyAsnAsnHisSerHisSer 285
QY      969  GAGGCGCTGTCGAGTGGCTGAGGACAAACCTTCAGAGAGCTGTGACCTCAGCCAG 1028
Db      286  GIUAlaLeuLeuAspGlyLeuAspGlyHisProSerHisGluLeuLeuThrLeuSerGln 305
QY      1029  AAAACAAAGCTTTTCTACTGAGTGGTGTGATGCTTTCTTCCATAGACCTCTGAT 1088
Db      306  LysThrLysSerPheThrArgValAlaValAspGlyAlaPhePheProHisGluProLeuAsp 325
QY      1089  CTATGTCCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACACAGAG 1148
Db      326  LeuLeuSerGlnLysAlaPheLysAlaLeuProSerLeuLeuGlyValAsnAsnHisGlu 345
QY      1149  TGTGGCTTCCTGCTGCTGATGAGAGGCTCCTGAGATCCCTCATGCTGCTCCACAACTCC 1208
Db      346  CysGlyPheLeuLeuProMetLysGlnAlaProGlnLeuLeuSerGlySerAsnLysSer 365
QY      1209  CTTCGCTCCATCTGATACAAACATCCTGACATCCCGCTCAGTATTGACCTTGTG 1268
Db      366  LeuAlaLeuHisLeuLeuLeuGlnAsnLeuLeuHisLeuProGlnLysLeuHisLeuVal 385
QY      1269  GCTAATGAAATCTTCCATGACAGCACTCCCTGATCAATTCGAGACAGTCTTTCTGAC 1328
Db      386  AlaAsnGlnLysPheHisAspLysHisSerLeuThrGlnLeuArgAspSerLeuLeuAsp 405
QY      1329  TTGCTTGAGATGATGTTCTTTGTGTCCTGACAGTGTACAGCTGATTCACAGAGAT 1388
Db      406  LeuLeuGlnLysPheValPhePheValAlaProAlaLeuLeuLeuHisLysArgLysHisArg 424
QY      1389  GCGGTGACACCTGTCTACTTCTATGAGTTTCGGACCCGCTCAGTGTGTAAGACACAG 1448
Db      424  -----
QY      1449  AAGCGGCTTTTGTCAAGCCGACACGCTGATGAGTCCGCTTGTGTGGTGGTGGCC 1508
Db      424  -----
QY      1509  TTCCTGAAGGGGACATTGTATGTTCCAGAGACCAAGAGAGAGAGAAATTACTGAGC 1568
Db      425  -----
QY      1569  CGAAGATGATGAAATGACTGGCTACTTGTGCTCGAAGCGGGAATCCATTAAGGAAAGAC 1628
Db      436  ArgLysMetMetLysLysTrpAlaThrPheAlaAspGlnGlyAsnProAsnGlyAsnAsp 455
QY      1629  CTGTCTCTGTGGCAGCTTATAATCTGACTGAGAGTACTTCAGCTGAGACTTGAACATG 1688
Db      456  LeuSerLeuTrpProAlaLysAsnLeuThrGlnGlnLysLeuLeuAspLeuAsnMet 475
QY      1689  AGCTTCGAGACAGACTCAAGAAAGACCGGGGTGATTTTGGACCAAGACCAATCCCTG 1748
Db      476  SerLeuGlnGlnAspLeuLeuLeuProArgValAlaPheTrpHisSerThrIleProLeu 495
QY      1749  ATCTGCTGCTGCTCGACATGCTCCAGAGTCTCTTCTCTTAATCTTCTGCTCTCTC 1808
Db      496  IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 515
QY      1809  CTCAGGCTTCTTTCTTTTCTTTGCTCTCT 1838
Db      516  LeuGlnProPhePhePheCysAlaPro 525

```

```

OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RX      PubMed=14702039; DOI=10.1038/ng1285;
RA      Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA      Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA      Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA      Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahata K.,
RA      Murakami K., Yasuda T., Iwayanagi T., Wagatsuma K., Shintani A.,
RA      Sudo H., Hosono T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA      Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa M., Yamazaki M.,
RA      Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA      Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA      Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA      Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA      Kanehori K., Takahashi-Fuji A., Hara H., Tanase T., Nomura Y.,
RA      Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA      Musuhino K., Yuki H., Oshino H., Sasaki N., Aotaka S.,
RA      Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA      Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA      Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA      Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA      Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA      Fujimori Y., Komiyama M., Tashiro A., Tanigami A., Fujimura T.,
RA      Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Okamoto S.,
RA      Kawabata A., Hikiji T., Kobatake N., Inagaki S., Ikema Y., Okamoto S.,
RA      Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA      Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA      Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA      Okumura K., Nagase T., Nomura N., Kikuchi H., Maehuo Y., Yamashita R.,
RA      Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT      "Complete sequencing and characterization of 21,243 full-length human
RT      cDNAs."
RL      Nat. Genet. 36:40-45(2004).
CC      -1 SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR      EMBL; AK090997; BAC03565.1; -.
DR      HSP; P12337.1K4Y.
DR      GO: GO:0016787, F:hydrolase activity, IEA.
DR      InterPro: IPR002018, Carboxylesterase.
DR      InterPro: IPR000379, Ser_ester.
DR      Pfam: PF00135, Coesterase; 1.
DR      PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR      PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW      Hydrolase.
SO      SEQUENCE. 469 AA; 52312 MW; 7A188F68C10A9080 CRC64;

Alignment Scores:
Pred. No.: 4,14e-172 Length: 469
Score: 2454.00 Matches: 467
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.57% Mismatches: 0
Query Match: 63.38% Indels: 0
DB: 2 Gaps: 0

US-10-023-515-1 (1-2158) x Q8NBC8 (1-469)
QY      432  ATGCTCAAGAGTCATTAACCCGAAATTTGAGTGTGAGAAGACTGCTTACTGAACATC 491
Db      1  MetLeuLysValHisLysTrpProLysPheGlyValSerGlnAspCysLeuThrLeuAsnIle 20
QY      492  TATGGCGCTGCCACGCGGATACAGAGCTCAAGCTCCCGCTTGAGTGGTCCCGAGA 551
Db      21  TyrAlaProAlaHisAlaAspPheThrGlySerLysLeuProValLeuValTrpPheProGly 40
QY      552  GGTGCTTCAAGACTGCTGCTCAGCTCCATCTTTGATGGTGGCTGGCTGCTGATGAG 611
Db      41  GlyAlaPheLysThrGlnGlySerAlaSerLysLysPheAspGlySerAlaLeuAlaLysGlu 60
QY      612  GACGTGCTGTTGGTGGTGTGCTCAGACCGGCTAGAGATATTGGTTCTTACACACATGG 671

```

```
Db      61 AspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThrTrp 80
Qy      672 GATGAGCAATGCTCCGGGGAACTGGGCTTCAAGAACAGATGGCTGTGCTCGGGGC 731
Db      81 AspGlnHisAlaProGlyAsnThrAlaPheLysAspGlnValAlaIleuSerTrpVal 100
Qy      722 CAGAGAACAATCAGAGTCTTCCGGTGGGGAGCCAGCTGTGATACCATCTTTGGCAGATCC 791
Db      101 GlnLysAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGlnSer 120
Qy      792 GCGGAGCCCAATAGTGTCTTATAGTCTTATAGTCTCCATGCGCAAGGCTTATTCAC 851
Db      121 AlaGlyAlaIleSerValSerSerIleuSerProMetAlaIleGlyLeuPheHis 140
Qy      882 AAGGCATCANTGAGAGAGTGGGGGCATCATCCCTTACTGAGAGCCCATATATATGAG 911
Db      141 LysAlaIleIleMetGlnSerIleValAlaIleIleProTyrLeuGlnLysAspTyrGln 160
Qy      912 AAGAGTGAAGACCTCAGGTGGTGGCAATTTCTGTGTAAACATGCTCAGACTGTAG 971
Db      161 LysSerGlnAspLeuGlnValValAlaHisPheCysGlyLysAsnAsnHisSerAspSerGln 180
Qy      972 GCGCTGTGAGGTGCTTGAAGACAAACCTTCGAAGAGCTGCTGACCTCAGCCAGAAA 1031
Db      181 AlaLeuLeuArgCysLeuAspGlyThrLysProSerLysGlnLeuLeuThrLeuSerGlnLys 200
Qy      1032 ACAAGTCTTTTACCTCGAGTGGTGTATGTGCTTTCTTCCATTAAGAGCTTATATCTTA 1091
Db      201 ThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGlnProLeuAspLeu 220
Qy      1092 TTGTCTCAGAAAGCAATTTAAACAATTCCTTCATCATCGAGTCAATTAACACAGAGTGT 1151
Db      221 LeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnSerHisGlnCys 240
Qy      1152 GCGTCTCTGCTGCTCCTATGAAAGAGGCTCCTGAGATCTCAGTGGCTCCAAAGTCCCTT 1211
Db      241 GlyPheLeuLeuProMetCysGlnLysAlaProGlnIleLeuSerGlySerAsnLysSerLeu 260
Qy      1212 GCGCTTCATCTGATACAAACATCTCGACATCCCGCTCAGTATTTGACCTTGTGGGT 1271
Db      261 AlaLeuHisLeuIleGlnAsnIleLeuHisIleProGlnIleuHisLeuValAla 280
Qy      1272 AATGAATCTTCATGACAGACAGCTCCCTGACCTGAAATCCGAGACAGTCTTGTGACTTG 1331
Db      281 AsnGlnTyrPheHisIleAspLysHisSerLeuThrGlnIleArgAspSerLeuAspLeu 300
Qy      1332 CTGGAAGATGTGTTCTTTGTGTGCTCCGTGACATGATCAGCTGATATCAGAGATGCT 1391
Db      301 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAspAla 320
Qy      1392 GGTGCACTGTCTTACTTCTATGAGTTTGGGACCCGGCTCAGTCTTTGAAGACAGCAAG 1451
Db      321 GlyAlaProValTyrPheTyrGlnPheArgHisArgProGlnCysPheGlnAspThrLys 340
Qy      1452 CCGGCTTTTGTGAAAGCCGACACAGCTGATGAAGTCCGCTTGTGTGGTGGTGGCTTC 1511
Db      341 ProAlaPheValLysAlaAspHisAlaAspGlnValArgPheValPheGlyIleAlaPhe 360
Qy      1512 CTGAAGGGGAGCATTTGATTATGTTCCGAAGAGCCACGAGAGAGAGAGTACTGAGCCGG 1571
Db      361 LeuLysGlyAspIleValIleMetPheGlnGlnGlyAlaThrGlnGlnGlnLysLeuSerArg 380
Qy      1572 AAGATGATGAATTAATCTGGGCTACCTTTGCTCGAACCGGGAATCTTAATGGAAACGACTTG 1631
Db      381 LysMetMetLysTyrTrpAlaThrPheAlaArgTyrGlnLysAsnProAsnGlnLysAsnAspLeu 400
Qy      1632 TCTCTGTGGCCGAGCTTATATCTGACTGAGAGTACCTCAGCTGAGACTTGAACATGAGC 1691
Db      401 SerLeuThrProAlaTyrAsnLeuThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSer 420
Qy      1692 CTCGAGCAGAGACTCAAGAACCCGGCGGTGAGATTTTGGACAGACCAATCCCGCTGATC 1751
Db      421 LeuGlyGlnArgLeuLysGlyIleProArgValGlnPheThrThrSerThrIleProLeuIle 440
```

```
Qy      1752 CTGTGCGCTCCGACATGCTCCACAGTCCCTTTTCTTCACTTCTCTCTCTC 1811
Db      441 LeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeuLeu 460
Qy      1812 CAGCTTTCTTTTCTTTTCTTTTGTGCTCCT 1838
Db      461 GlnProPhePhePhePheCysAlaPro 469

RESULT 4
06AM47
ID 06AM47 PRELIMINARY; PRT; 575 AA.
AC 06AM47;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxin;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN (1)
RP SEQUENCE FROM N.A.
RA Miyazaki M., Yamashita T., Taira H., Suzuki A.;
RT "cauxin family protein.";
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: belongs to the type-B carboxylesterase/lipase family.
EMBL; AB186392; BAD35015.1; -.
GO GO:0003824; F:catalytic activity; IBA.
DR InterPro; IPR002018; CarboxesteraseB.
DR InterPro; IPR000379; Ser ester.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 575 AA; 63620 MW; 1005C35B82E1183D CRC64;

Alignment Scores:
Pred. No.: 4,93e-157 Length: 575
Score: 2250.00 Matches: 417
Percent Similarity: 86.75% Conservative: 61
Best Local Similarity: 75.66% Mismatches: 73
Query Match: 58.11% Indels: 0
DB: Gaps: 0

US-10-023-515-1 (1-2158) x 06AM47 (1-575)
Qy      186 GGGCTTCTGCTGAAGGGCCACAGAGAAACACAGCTGGATGATTCAAGGCAAGCA 245
Db      25 GlyProAlaAlaAspAlaProValArgSerThrArgLeuGlyTrpValArgGlyLysGln 44
Qy      246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTGAAGTCCCTTTGCTGCT 305
Db      45 AlaThrValLeuGlySerThrMetProValAsnValPheLeuGlyIleProPheAlaAla 64
Qy      306 CCCCCGCTGGAGTCCCTCGATTTACGAACCCGAGCCTGCATCCGCCCTGGGATAACTTG 365
Db      65 ProProLeuGlyProLeuArgPheLysArgProLysProAlaLeuLeuThrAsnAspSer 84
Qy      366 CGAGAAGCACTCTTACCTTAATTTGTGCTCCGAAGTCAAGATGAGTGGCTCTTAGAT 425
Db      85 ArgAspAlaThrSerTyrProLysLeuCysLeuGlnAsnSerValTrpLeuLeuSerAsp 104
Qy      426 CAACATGCTCAAGTCAATTAACCGAAATTCGAGGTGTCAAGAGTGCCTTACCTG 485
Db      105 GlnHisPheLeuLysValHisTyrProAsnLeuGlnValSerGlnAspCysLeuTyrLeu 124
Qy      486 AACATCTATGGCTGGCCCGCCAGCGCATACAGGCTCCAGCTCCCGTCTTGTTGTGCTTC 545
Db      125 AsnIleTyrAlaProAlaHisAlaAsnThrGlySerLysLeuProValMetValTrpPhe 144
Qy      546 CCAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAATGGGTGCGCCCTGCTGCC 605
```

Db 145 ProGlyGlyAlaPheGluThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa 162

Qy 606 TATGAGAGCTGCTGGTGTGTGCTGCTGACAGTACCGGCTAGGAAATATGATTTCTACAC 665

Db 165 TGTGTAAPValLeuIleValThrThrGlnTyrArgLeuGlyIlePheGlyPhePheLys 185

Qy 666 ACATGGAGATCAGCATGCTCCGGGGAACTGGGCTTCAAGACACAGGTGGCTGTGCTCC 725

Db 185 ThrGlyAPGlnIleAlaProGlyValThrAlaPheLeuAspGlnLeuAlaAlaLeuThr 204

Qy 726 TGGGTCCAGAAAGACATGAGTTCTTCGGTGGGGACCCAGACTCTGTGACATCTTGGC 785

Db 205 TrpValGlnGluAenIleGluPhePheGlyGlyAProPheIleSerValThrIlePheGly 224

Qy 786 GAGTCCGGGGAGCCATAAGTGTTCATAGTTTATACTGTCTCCAGTGGCCAAAGCTTA 845

Db 225 GluSerAlaGlyAlaIleSerValSerGlyLeuValLeuSerProMetAlaSerGlyLeu 244

Qy 846 TTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGAT 905

Db 245 PheIleLysAlaIleMetGluSerGlyValAlaIleIleProPheLeuArgAlaProAsp 264

Qy 906 TATGAGAAAGTGAAGACTGTCAGCTGTGTTCACATTTCTGTGTATCAATGCGTCAGAC 965

Db 265 AspGlnTyrAsnGluAProLeuGlnValIleAlaArgIleCysGlyCysAsnValSerAsp 284

Qy 966 TCTGAGAGCCCTGCGAGGTGCTCGAGAGCAAAACCCGCCAAGAGAGCTGTGACCCCTGAC 1022

Db 285 SerValAlaLeuLeuGlnCysLeuArgAlaLysSerSerGluIleuLeuAspIleAsn 304

Qy 1026 CAGAAACAAAGTCTTTCACATGAGGTGTGATGATGTCTTCTTCTCTATAGACCTTA 1085

Db 305 LysAlaThrLysSerPheThrArgValIleAspGlyPhePhePheProAspGluProLeu 324

Qy 1086 GATCTATGTCCTCAGAAAGCATTTAAAGCAATTCCTCATCATGCGAGTCAATACCAAC 1145

Db 325 AspLeuLeuThrGluLysThrPheAsnSerIleProSerValIleGlyValAsnAsnIle 344

Qy 1146 GAGTGTGCTCTCTGCTGCTATGAAGAAGGCTCCGATGATCTCGAGTGGCTCCAAAG 1205

Db 345 GluCysGlyPheLeuLeuProMetLysGluPheProGlnIleLeuGlyGlySerAsnLys 364

Qy 1206 TTCCTTGACCTCATGTGATTAACAAACATCTCGACATCCGCGCTCAAGTATTTGACCTT 1265

Db 365 SerLeuAlaLeuAsnIleuLeuIleIleArgValLeuAsnIleProAsnGlnTyrLeuTyrLeu 384

Qy 1266 GTGGCTAATGAATACTTCCATGACAAAGACTCCCTGACGTAAATCCGAGACAGTCTTGG 1325

Db 385 ValAlaAspGlnTyrPheTyrAsnLysIleSerProValGluIleArgAspSerPheLeu 404

Qy 1326 GACTTGCTTGAGAGTGTGTTCTTGTGCTCCCTGACATGATCAGACTCGATATCAGAGA 1385

Db 405 AspLeuLeuLeuGlyAspValPhePheValValIleProGlyValIleThrAlaArgTyrIleArg 424

Qy 1386 GATGTGTGTCACCTGTACTTCTATAGTTCGGGACCGGCGCTCAAGCTTTGAAGAC 1445

Db 425 AspAlaGlyAlaProValTyrPheTyrGluPheGlnIleSerProGluCysValLysAsp 444

Qy 1446 ACGAAGCCGGCTTTGTCAAAAGCGGACCAAGCTATGAAGACTCCGTTGTGTTCGGGTGT 1505

Db 445 ThrArgProAlaPheValValAlaAspIleSerAspGlnIleArgPheValPheGlyGly 464

Qy 1506 GCCTTCCGAGAGGGGACATGTTATGTTGGAAGAGCAACCGAGGAGGAGAAATTATCTG 1565

Db 465 AlaPheLeuLysGlyAsnIleValMetPheGlnIleValThrGlnGlnGluLysValLeuLeu 484

Qy 1566 AGCCGAAAGATGATGAATACTGGGCTACCTTGTGCTGAAACCGGGAATCTTAATGGGAAC 1625

Db 485 SerArgLysMetMetArgTyrTTPAlaAsnPheAlaArgThrGlyAProAsnGlyGlu 504

Qy 1626 GACTGTCTCTGTGGCGCATTAATCACTGACGAGAGTACTCCAGTGCATTGGAAC 1685

Db 505 GlyLeuProLeuTriProAlaIarYserGlnSerGlnGlnIntYLeuYbLeuAspLeuAsn 524

QY 1686 ATAGCCTCGGACAGAGACTCAAGAACCGGCGGTGATCTTTTGACACGACGATCC 174

Db 525 IIsErVaIGlYgInLeuYseGlnInuIuValGluHetpSerApHrLeuPro 544

QY 1746 CTGATCTCTTCGCTCCGACATGCTCCACATGCTCTTTCTTCCTTAACTTTCCTCT 180

Db 545 LeuIleMeSerMeISeSerThralaProProGlyProProValProLeuLeuSerLeuSer 564

QY 1806 CTCCTTCAGCCTTTCTTTTCTTTTGGGCTCT 1838

Db 565 ValLeuLeuProPheLeuPheSeSerAlaPro 575

RESULT 5
06AW46

ID	Q6AW46	PRELIMINARY;	PRT;	575	AA
AC	Q6AW46.				

DT 25-OCT-2004 (TREMBlrel. 28, Created
DT 25-OCT-2004 (TREMBlrel. 28, Created

DT 25-OCT-2004 (Tremblay, 28, Last annotation update)

GN	Name=cauxin;
MS	Mus musculus (M-mus)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorhina; Carnivora

OX NCBI_TaxID=10090
PN [1]

RP SEQUENCE FROM N.A.
85

RT	cauxin family protein.;	Yasuda H.; Suzuki A.;
RT	"cauxin family protein.;"	Suzuki A.;

-1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family

GO; GO:0003824; F:catalytic activity; IEA

InterPro; IPR000379; Ser esters.

PROSITE; PS00122; CARBOXYLESTER

Hydrolase.

04108 MM; 3EB9D85981D9DE0A CRC64;

Assignment scores:	
Pred. No.:	1.45e-145
Length:	575

Percent Similarity:	2094.50	Matches:	390
	80.67%	Conservative:	69

Local similarity:	68.54%	Mismatches:	99
Query Match:	54.09%	Indels:	11

2 Gaps: 1

03-10-023-515-1 (1-2158) X Q6AW46 (1-575)

129 |GGIGCTTTTCTGATTCACAGCCCTGTTGGACACAGACAGTGGGAAACTGGG 188

1 / rrp1pnehg1yala1e1le1e1u-----Gly 25

189 CTTCTGTGAGGGCCACAGAGGAACACACGCTGGATGATTCAGGCCAAGCAAGTC 248

26 ServaAlThrGluIuProHisArgTyrThrLysLeuGlyTyrPvalGlnGlyLysGlnAla 45

249 ACTGTGCTGGGAGCCCTGTGCTGTGACGTGTTCTCCGAGTCCCCCTTGCTGCTCCC 308

46 ThrValLeuGlyArgLeuGluProValAsnValPheLeuGlyIleProPheAlaAlaPro 65

309 CCGCTGGATCCCTGCGATTACGAACCCGACCTGCATCGCCCTGGATACTTGGCA 368

66 ProLeuGlyProLeuArgPheSerLysProGlnProProIleProTyrPaspAsnLeuArg 85

369 GAAGCCACCTCCTACCTAATTGTGCCCTCAGAACTCAGAGTGGCTGCTCTTAGATCAA 428

86 GluAlaThrAlaTyrProAsnLeuCysPheGlnAsnLeuGluTrpLeuPheIleTyrGln 105

QY	429	CACATGCTCAAGGTGCATTACCCGAAATTCGAGGTGCAGAAAGTGCCTTACTGAC	488
Db	106	AsnLeuLeuLysValSerTyrPro1LeuL1yMetSerGlnAapCybLeuTyrLeuAsn	125
QY	489	ATCTATGCGCTGCGCCACGCGCGATACAGAGCTCCAAAGCTCCCGCTTGATGGTCCCA	548
Db	126	IleTyrAlaProCysH1eValAlaAsnAsnD1ySerSerLeuProValMetAlrTpr1ePro	145
QY	549	GAGAGTGCCTTCAAGATGCGCTCAGCCTTCATCTTTGATGGGTCCGCGCTGGCTGAT	608
Db	146	G1yG1yG1yPhcG1uThrG1ySerAlaSer1ePhcAapG1ySerAla1eAla1yAlrYx	165
QY	609	GAGACCTGCTGGTGTGGTGGTGCACAGTCCGCGCTAAGAAATTTGGTTTTCACACCA	668
Db	166	G1uAapVal1eAlaValThr1eG1uTyrAlaG1yG1y1ePhcG1yPhcPhcThrThx	185
QY	669	TGGATATCAGATGCTCCGCGGGAATCGAGGCTTCAAAGACACAGTGGCTGTGCTGG	728
Db	186	G1uAapG1uH1eValAlaProG1yAsnTTrAlaPhcTTrAapG1uAla1eAla1eLeuTTr	205
QY	729	GTCCAGAAAGACATCGAGTCTTTCGGTGGGAGCCCGACCTGTGCATCTTTGGCGAG	788
Db	206	ValArG1uAen1Ie1yTyrPhcG1yG1yAlaAapProAapSerValThr1ePhcG1yAa	225
QY	789	TCCCGGGAGCGCTAAGTGTTC1ATGCTTATATCTGTCTCCCATGCGCAAGGCTTATC	848
Db	226	Ser1aG1yAla1eSer1eSerSerLeu1eLeuSerProLeuSerAlaAapLeuPhc	245
QY	849	CACAAAGCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGGCCCATGATAT	908
Db	246	H1sTrAla1IeMetG1uSerG1yAlaAla1e1eProSerLeuLysSerSerAapAa	265
QY	909	GAGAAAGTGAAGCATCTGACAGTGGTGGTGCATCTTGTGGTAAACATGCGCTGACTT	968
Db	266	AspLeuLysH1eAapLeuG1uValAlaAlaAaH1yCybAapCybAaH1ySerAapSer	285
QY	969	GAGCGCTGCTGAGGTGCCTGAAGCAAAACCTTCAAAGATGCTGACCTCAGCCAG	1028
Db	286	LysAlaLeuLeuLysCybLeuAlaG1yLysSerSerLeuG1uMetSerLeuSerG1u	305
QY	1029	AAAACAAAAGTCTTCACTCGAGTGGTATGATGGTCTTCTTCCATAGACCTCTGAT	1088
Db	306	LysAlaLysSerPhcThrAlaValAlaAapLysSerPhcPhcSerG1uG1uProLeuG1u	325
QY	1089	CTATTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCCATCAGCGAGTCAATTAACAGAG	1148
Db	326	LeuLeuSerG1uLysThrLeuLysAla1eAlaProSer1e1eG1yAlaAsnAaH1uG1u	345
QY	1149	TGTGGCTTCTGCTGCTCATGAAGAGGCTCTGAGATCTCACTGCTCCAAAGTCC	1208
Db	346	CysG1yTyr1eLeuProValAlaArgAapThrProG1u1eLeuG1ySerAlaH1uSer	365
QY	1209	CTTGGCCCTCATGATACAAACATCCTGACATCCGCGCTCAAGTATTCACCTGTG	1268
Db	366	ThrAlaLeuThrLeu1eH1eThrLeuLeuH1eVal1eProH1uG1uH1yLr1eAla1	385
QY	1269	GCTATGATATCTTCCATGACAAAGACCTCCCTGATCAATTCGAGACCTTGTGAC	1328
Db	386	ThirLysG1uTyrPhcH1eSg1yLysH1eSerProThrAap1eAlaGApThrLeuLeuAap	405
QY	1329	TTGCTTGAGATGTGTCTTGTGTGCTCCTGACATGATCAGAGCTCGATATCAGAGAT	1388
Db	406	LeuPhcG1yAapValPhcPhcAlaValAlaProG1yLeuVal1ThrAlaArgTyrH1eAap	425
QY	1389	GCTGTGACCTGTCTACTTCTATAGAGTTTGGGACCGGCGCTCAAGTCTTTGAAGACAG	1448
Db	426	SerG1yG1yProValTyrPhcTyrG1uPhcG1uH1eAapProH1eCybPhcG1uAapSer	445
QY	1449	AAGCGCGCTTTTGTCAAGCGGACACACCTGATGAAGTCCGCTTGTGTGGTGGTCC	1508
Db	446	ArgProAlaPhcValLysAlaAapH1eThrAapG1u1eAlaGPhcValPhcG1yG1yPro	465
QY	1509	TTCTCGAAGGGGACATTGTATTATGTTTCAGAGGACACGAGAGAGAAAGTATCTGAC	1568

[illegible]


```

Db      45 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTyrAlaIa 64
Qy      306 CCCCGCTGGAGTCCCTGGATTTCAGAACCCGCGAGCTGCATCGCCCTGGAGTAATCTTG 365
Db      65 ProPheLeuGlyIleProLeuArgPheLeuGlyIleProValAlaLeuProGlyIleAsn 84
Qy      366 CGAAGAGCCCTCTCCATTAATTTGGCTCCGAGAGCTGAGAGTGGCTGCTTGGAT 425
Db      85 ArgAsnAlaThrSerTyrProLeuLeuGlySerPheGlnAblLeuGlnIlePheValSerTyr 104
Qy      426 CAACAGATGCTCAAGATGATTAACCGAAATTCGAGTGTCAAGAACTGCTTACCTG 485
Db      105 GlnHisValLeuValArgTyrProLeuGlnIleAsnSerGlnAblProGlyLeuIleu 124
Qy      486 AACATCTATGGCGCTCCGAGCGAGTACAGAGCTCAAGCTCCCGCTTGGTGTGCTTC 545
Db      125 AsnIleTyrAlaProAlaIleAlaAspAsnIleSerAsnLeuProValIleValTyrPhe 144
Qy      546 CCAGAGAGTGGCTTCAGAGCTGGCTCAGCTTCATCTTTGATGGGTCCGCTGGCTGCC 605
Db      145 ProGlyValAlaPheLeuMetGlySerAlaSerSerPheAspGlySerAlaLeuAlaIa 164
Qy      606 TATGAGAGCTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
Db      165 TyrGlnAspValLeuIleValThrThrGlnTyrArgLeuGlyIlePheGlyPhePheAsp 184
Qy      666 ACATGGATCAGCATGCTCCGAGGAGCTGGCTTCAGAGACAGAGTGGCTGTGTGTCC 725
Db      185 ThrGlyAspGlnHisAlaArgGlyAsnTyrAlaLeuLeuAspGlnAlaAlaLeuThr 204
Qy      726 TGGGTCCAGAGAAACATCGAGTCTTGGTGGGAGACCCGAGCTGTGACCATCTTTGGC 785
Db      205 TrpValArgAspAsnIleGlnPhePheGlyIleAspProArgSerAlaThrIlePheGly 224
Qy      786 GAGTCCGCGGAGCCATAGTCTTCTAGCTTATCTGCTGTGTGTGTGTGTGTGTGTGTGT 845
Db      225 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAsnGlyLeu 244
Qy      846 TTCACAAAGCCATCATGAGAGTGGGTGGCTCATCATCTTACCTG--GAGGCCCAT 902
Db      245 PheHisValAlaIleMetGlnSerGlyValAlaIleLeuProLeuMetArgProPro 264
Qy      903 GATTATGAGAGAGTGAAGCTGAGAGTGTGTGACATTTCTGTGTGTGTGTGTGTGTGTGT 962
Db      265 GlyAspGlnArgGlyAspAsnLeuGlnIleValIleAlaArgIleCysHisAlaSer 284
Qy      963 GACTCTGAGGCTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
Db      285 AspSerAlaAlaLeuLeuGlnCysLeuArgAlaIleProSerGlnGlnLeuMetAspIle 304
Qy      1023 ACCCAAGAAACAAAGCTTTCATCTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082
Db      305 SerIleValLeuThrPheSerIleProValIleAspAspPhePheProAspIlePro 324
Qy      1083 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATGAGTCAATTAAC 1142
Db      325 ValAlaLeuThrGlnIleGlnAlaPheAsnSerValProSerIleIleGlyValAsnAsn 344
Qy      1143 CACGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1202
Db      345 HisGlnCysAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyIleSerAsn 363
Qy      1203 AAGTCCCTTGGCTTCATCTGATCAAAACATCTTCATCTCCGCTCAGTATTTGAC 1262
Db      364 ArgSerIleValAlaLeuTyrLeuValHisThrPheLeuAsnIleProThrGlnTyrLeuHis 383
Qy      1263 CTGTGGCTATGAAATTAATCTCAGACAGACATCTCCGACTGAATCCAGACAGTCTT 1322
Db      384 LeuValAlaAspHisTyrPheTyrAsnIleHisSerProValGlnIleArgAspSerPhe 403
Qy      1323 CTGAGACTGTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1382
Db      404 LeuAspLeuLeuGlnIleAspValIleLeuPheValValProGlyValAlaThrAlaArgTyrHis 423

Qy      1383 AGAGATGCTGTGACACTGTGTCTTATGATGTTTCCGACCGGCTCAGTCTTTGAA 1442
Db      424 ArgAspAlaGlyAlaProValIleTyrPheTyrGlnPheGlnHisProProGlnCysLeuAsn 443
Qy      1443 GACACGAGCCGCGCTTTTGTCAAGCCGACACAGCTGATGAAAGTCCGCTTGTGTGCGT 1502
Db      444 AspThrArgProAlaPheValIleAspHisSerAspGlnIleArgPheValPheGly 463
Qy      1503 GGTGCTTCTGTAAGGGGGGACATTTGTTATTTTTCGAAGACCGAGAGAGAGAGTTA 1562
Db      464 GlyAlaPheLeuValGlyAspIleValMetPheGlnGlyAlaThrGlnGlnGlyLeu 483
Qy      1563 CTGACCCGGAAGATGATGAATTAATCTGAGTACTTTGTCTCAACCGGGAATCTTAAGGG 1622
Db      484 LeuSerArgIleMetMetCysTyrTrpAlaAsnHisAlaGlnTrpGlyAspProAsnGly 503
Qy      1623 AACGACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1682
Db      504 GlnGlyValIleProLeuTyrProAlaTyrThrGlnSerGlnGlnIleTyrLeuValLeuAspLeu 523
Qy      1683 AACATGACCTCGGACAGAGACTCAAGAAACCGGAGGATTTTGGACACAGACATC 1742
Db      524 SerValSerValGlnGlnIleValLeuLeuGlnGlnIleValGlnPheTrpMetAsnThrIle 543

RESULT 7
Q95KH3 PRELIMINARY; PRT; 361 AA.
AC 095KH3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey) .
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euleleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_Taxid=9541;
RN [1]
RP
RC SEQUENCE FROM N.A.
RT TISSUE=Temporal lobe right;
RA Oesda N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.,
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AB060873; BAB46884.1; -.
DR HSSP: P12337; 1K4Y.
DR InterPro: IPR002018; Carboxylesterase.
DR Pfam: PF00135; Coesterase; 1.
KW Hypothetical protein.
SQ
SEQUENCE 361 AA; 41032 MW; AC62AB6E78C5F1A1 CRC64;

Alignment Scores:
Pred. No.: 6,37e-116 Length: 361
Score: 1691.50 Matches: 327
Percent Similarity: 96.23% Conservative: 5
Best Local Similarity: 94.78% Mismatches: 12
Query Match: 43.69% Indels: 1
DB: 2 Gaps: 1

US-10-023-515-1 (1-2158) x Q95KH3 (1-361)
Qy      819 ATACTGTCTCCATGGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGTGGCC 878
Db      17 IleLeuSerPrometAlaGlnGlyLeuPheHisValAlaIleMetGlnSerGlyValAla 36
Qy      879 ATCATCCCTTACCGGAGCCCATGATTTATGAGAAAGTGAAGACTTCGACAGTGGTTGCA 938
Db      37 IleIleProTyrLeuValAlaTyrAspTyrGlnIleYsSerGlnPheLeuGlnValAlaIa 56
Qy      939 CATTTCTGTGTAACAATGCTCAGACTGTGAGGCTCGTGGTGGTGGTGGTGGTGGTGGTGG 998
Db      57 HisPheCysGlnAsnAlaAsnAlaSerAspSerGlnAspLeuLeuArgCysLeuArgThrLys 76

```



```

OY 999 CCTCCAGAGAGCTGCTGACCTCCAGCCGAGAAACAAAGCTTTTCTGAGTGGTGT 1058
DB 77 SerSerLySGluLeuLeuThreSerGlnLysAlaIleSerPheThrArgValValAsp 96
OY 1059 GGTGCTTTCTTCTTAATGAGCCTTGAATCTATTGTCTCAGAAAGCATTTTAAGCAAT 1118
DB 97 GAlaIaPhePheProAsnGluProLeuGlnLysLeuSerGlnLysAlaIaPheValaIle 116
OY 1119 CCTTCATCATCTGAGAGTCAATTAACCAAGGTGGCTTCTGCTGCTCTATGAGAGAGCT 1178
DB 117 ProSerIleIleGlyValAsnAsnHisGluCysGlyPheLeuLeuProMet---GluAla 135
OY 1179 CCTGAGATCTCAGAGGGCTCCAGAAAGTCCCTTGGCCCTCATCTGATTAACAAACATCTCG 1238
DB 136 ProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnHisIleLeu 155
OY 1239 CACATCCCGCTCAGATATTGACCTTGTGCTAAATGAATATTCTCATGACAGCACTCC 1298
DB 156 HisIleProProGlnTyrLeuHisLeuValAlaAsnGluTyrPheHisMetLysHisSer 175
OY 1299 CTGATCGAAATCCGAGACAGTCTTCTGACCTTGTGAGATGTGTCTTTGTGCTCCT 1358
DB 176 LeuThrGlnIleArgAspGlyLeuLeuMetLeuGlnLysAspValPhePheValValPro 195
OY 1359 GCATGATCAGACGCTGATATCAGACAGAGATGCTGGGACCTGCTACTTCTATGAGTTT 1418
DB 196 AlaLeuIleHisAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPhe 215
OY 1419 CGGACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAGCCGACCAAGCT 1478
DB 216 ArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHisAla 235
OY 1479 GATGAAGTCCGCTTGTGTGCTGCTGCTGCTTCTGAGGGGACATGTTATGTTGAA 1538
DB 236 AspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlu 255
OY 1539 GGAGCGAGGAGGAGAGAGTACTGAGCGCGGAGAGATGATGAATPCTGGGCTACCTT 1598
DB 256 GAlaIaThrGlnGlnGlnLysLeuSerArgLysMetMetLysTyrTrpAlaThrPhe 275
OY 1599 GCTCGAACCGGAGATCTTAATGGAACGACCTGCTCTGTGTGCGCAGCTTAATCTGACT 1658
DB 276 AlaArgThrGlyAsnProAsnGlnLysAspLeuProLeuTrpProAlaTyrHisLeuThr 295
OY 1659 GAGCAGTACTCTCAGCTGACTTGAACATGAGCTTCCGACAGACAGACTCAAGACCGCG 1718
DB 296 GlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGluProArg 315
OY 1719 GTGATTTTGAACAGACACCATCCCTGATCTGTCTGCTCCGACATGCTCCAGACT 1778
DB 316 ValGlnPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspThrLeuHisSer 335
OY 1779 CCTCTTTCTCTTAACCTTCTCTCTCTCTCCAGACCTTTCTTTTCTTTGTGCTCCT 1838
DB 336 ProLeuSerPheLeuIlePheLeuSerLeuLeuGlnProPhePheSerCysValPro 355
OY 1839 TGAGAGATTATCTTT 1853
DB 356 TrpGluValIlePhe 360

```

```

RESULT 8
SASB_ANAPL STANDARD; PRT; 557 AA.
AC Q04791;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DB Pctyl acyl-CoA hydrolase precursor, medium chain (EC 3.1.2.-)
DB (Thioesterase B).
DB Anas platyrhynchos (Domestic duck).
DB Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.

```

```

OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-65.
RC TISSUE=utrocytic gland;
RX MEDLINE=9330823; PubMed=8314791;
RA Hwang C.-S., Kolatukudy P.E.;
RT "Molecular cloning and sequencing of thioesterase B cDNA and
RT stimulation of expression of the thioesterase B gene associated with
RT hormonal induction of peroxisome proliferation.";
RL J. Biol. Chem. 268:14278-14284(1993).
CC -1- FUNCTION: Fatty acid biosynthesis chain termination and release of
CC the free fatty acid product is achieved by hydrolysis of the thio
CC ester by a thioesterase. This thioesterase may be associated with
CC peroxisome proliferation and may play a role in the production of
CC 3-hydroxy fatty acid diester phenones.
CC -1- TISSUE SPECIFICITY: Highest levels in utrocytic gland, much lower
CC in liver and kidney.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC
DR EMBL; L05493; AAA49223.1; -.
DR PIR; A47162; A47162.
DR HSP; 07540; 1K4Y.
DR InterPro; IPR002018; CarboxylesteraseB.
DR InterPro; IPR00379; Ser esterase.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR KW Direct protein sequencing; Fatty acid biosynthesis; Hydrolyase; Signal.
FT SIGNAL 1 25
FT CHAIN 26 557 Fatty acyl-CoA hydrolase.
FT ACT_SITE 227 227 Acyl-ester intermediate (By similarity).
FT ACT_SITE 345 345 Charge relay system (By similarity).
FT ACT_SITE 460 460 Charge relay system (By similarity).
FT DISULFID 93 122 By similarity.
FT CARBOHYD 476 476 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 557 AA; 61637 MW; 03E35D90A037E6B0 CRC64;

```

Alignment Scores:

```

Pred. No.: 7,146-87 Length: 557
Score: 1298.50 Matches: 265
Percent Similarity: 62.77% Conservative: 79
Best Local Similarity: 48.36% Mismatches: 177
Query Match: 33.54% Indels: 27
DB: 1 Gaps: 8

```

US-10-023-515-1 (1-2158) x SASB_ANAPL (1-557)

```

OY 133 ACTGGGCTTCTGCTGAGGCGCAAGAGAAACAGGCTGGATTCAGGGCAAG 242
DB 23 ThrGlyGlnLysAlaGlnGlnProGluValAlaThrAsnTyrGlySerValArgGlyTyr 42
OY 243 CAAGTCACCTGCTGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
DB 43 GlnValLysValAlaAsnAlaGlnLysGlnValAlaThrValPheLeuGlyLeuProPheAla 62
OY 303 GCTCCCGCTGGAGATCCCTGGATTAAGAACCGGACGCTGATCGCTGGATTAAC 362
DB 63 LysProProValGlyProLeuArgPheSerGlnProGlnProGlnProGlnProGlnProGln 82
OY 363 TTGGAGAGAGCACTCTTACCTTAATTTGGCTCCCGAAGTCAAGATGGTGTCTTCA 422
DB 83 ValArgAspAlaAlaSerTyrProMetCysLeuGlnAspLys-----ValLeu 99
OY 423 GATCAACATGATCTC-----AAGGTGATTACCCGAAATTCGAGTGTCA 467

```

```

Db      100 glyInIyLeuSeXAspAlaIleThraenAArgLySGluYValArgLeuGlnIleSer 119
Qy      468 GAAGATGCGCTCTACCTGAAACATCTATGGCGCCCGCCAGATACAGGCTCCAGATC 527
Db      120 GlnuApCyLeuIyLeuEnuValIyThrProValSerThrgluGlnGluYLeu 139
Qy      528 CCCGCTTGCTGTGGTCCCAAGAGCTCTTCAAGACTGGCTCGACCTCCATCTTGAT 587
Db      140 ProValPheValTrpIleHieGlyGlyLeuValSerGlyAlaAlaIleSerTyrAsp 159
Qy      588 GGGTCCGCGCTGGCTATGAGAGACTGCTGCTGTGGTCTGCTCAGTACCGGCTAGA 647
Db      160 GlySerAlaLeuAlaAlaPheAspAnValValValAlaTrpIleGlnIyArgLeuGly 179
Qy      648 ATATTGGTTCTTCCACCAATGGGATGATGATGCTCCGGGAACTGGGCTTCAAGAC 707
Db      180 IleAlaGlyTrpPheSerThrGlyAspIySerHieAlaArgGlyAnTrpGlyTrpLeuAsp 199
Qy      708 CAGGTGGCTGCTCTGCTCTGGGTCGAGAAACATGATCTTCCGTGGGAGCCCGAC 767
Db      200 GlnValAlaAlaLeuGlnTrpIleGlnGluAnIleIleHiePheAArgLyAspProGly 219
Qy      768 TCTGTGACCACTTTGGCAGATCCGCGGAGCCATAGTGTCTTACCTTATCTGCT 827
Db      220 SerValThrIlePheGlyGlySerAlaGlyValSerValSerAlaLeuValLeuSer 239
Qy      828 CCCATGGCCCAAGGCTTATTCACAAAGCATATGAGAGTGGGCTGCGCATC--ATC 884
Db      240 ProLeuAlaIySGlyLeuPheHieIyAlaIleSerGluSerGlyThraIaValArgIle 259
Qy      885 CCTTACCTGGAGGCGCATGATTATGAGAAGTGAAGACTGCAAGTGGTGGACATTTTC 944
Db      260 LeuPheThrIu-----GlnProGluGlnGlnAlaGlnIyAlaIleAlaAla 275
Qy      945 TGTGTGAACAATGCTGAGACTGAGGCTGAGGCTGAGGCTGAGGACAAACCTCC 1004
Db      276 AlaGlyCyneGluIySerSerSerAlaAlaLeuValGluCySerLeuArgIyThrglu 295
Qy      1005 AAGGAGCTG-----CTGACCCCTCAGCCGCAAAACAAAGCTTTTCACTGAGTGGTAT 1058
Db      296 AlaGlnMeGluGlnIleThrLeuIySwetProPometPheIleSerAlaSerLeuAsp 315
Qy      1059 GGTGCTTTCTTCTCTATGAGCTCTTATGATCTTATGCTCAGAAAGCATTTAAACAT 1118
Db      316 GlyValPhePheProIySerProAArgGlnLeuLeuSerGlnIyValIleAnIaVal 335
Qy      1119 CCTTCATATATGGAGTCAATAACACAGATGCTGCTCTGCTCT--ATGAAGAG 1175
Db      336 ProTyrIleIleGlyValaAnAnCySGluPheGlyTrpIleLeuProArgMetLeuYs 355
Qy      1176 GCTCCTGAGATCCTGAGTGGCTCCAAGATCCCTTGGCTCATCTCATCAAAACATC 1235
Db      356 PheProGluPheThrgluGlyLeuGluIyAspValAlaArgGlnValLeuGlnIleSerThr 375
Qy      1236 CTGCACATC-----CCGCTCAGATTATTCACCTTGTGGTGAATGAATAC 1280
Db      376 LeuAlaLeuSerPheIyGlyAlaProSerAspIleValaPheLeuValTyrAsnGluTr 395
Qy      1281 TTCATGACAAAGCATCCCTGATGTAATCCGAGACATCTTTGGAGCTTCTGGAGAT 1340
Db      396 IleGlyValAlaGluAnArgAlaGlnValaIyArgAspGlyLeuLeuAspSerIleAlaAsp 415
Qy      1341 GTGTTCTTTGTGCTCCCTGACATGATCAGCTGATGATCAGAGATGCTGGTGGACCT 1400
Db      416 ProLeuPheValPheSerAlaValaGluValaIaArgHieShIaArgAspAlaGlyAspPro 435
Qy      1401 GTCTACTCTATAGTTTGGGACCGGCTCAGTGGCTTTGAAGACACAGAAAGCCGCTTT 1460
Db      436 ValIyTrpPheTrpGluPheGlnHieArgProSerSerAlaAlaGlyValaIyProGluPhe 455
Qy      1461 GTCAAGCGGACGACGATGATGAGTCCGCTTGTGCTGGTGGTGGCTTCTGGAAGGG 1520
Db      456 ValIyValaIaPheHieAlaIaPheGlnIleAlaPheValPheGlyIySProPheLeuAlaGly 475

```

```

Qy      1521 GACATTGTTATGTTCCAAAGGACACAGAGAGAGAAATTATCTGAGCCGGAAGATGATG 1580
Db      476 Asn-----AlaThrgluGlnGluAlaIyLeuLeuSerArgThraValMet 489
Qy      1581 AATATCTGGGCTACCTTTCTCTGAAACCGGAAATCTTAATGGAGACACTGCTCTGTGG 1640
Db      490 LyTrpTyrTrpThraAnPheAlaArgAnGlyAsnProAnGlyGluGlyLeuValHieTrp 509
Qy      1641 CCAGCTTATATCTGACATGATGATCCGAGTCCGAGCTTGAAACATGAGCTCGGACAG 1700
Db      510 ProGlnIyTrpAspMetAspGluArgIyLeuGlnIleAspLeuThrgluYAlaAlaIyS 529
Qy      1701 AAGCTCAAAAGAACCGGCTGATTTTGAACAGACATCCCTGATCTGCTGACC 1760
Db      530 LyLeuIySGluArgLyMeGluPheThrgluMetGlnLeuThrgluGlnIleMetSerAsp 549
Qy      1761 TCCGACATGCTCCACAGTCTCTT 1784
Db      550 ArgArgArgLyShIaThraPheLeu 557

```

```

RESULT 9
ID      Q91WG0
AC      Q91WG0;
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT      25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE      Carboxylesterase 2 (Carboxylesterase M13 precursor).
GN      Name=Ces2; Synonyms=ces2A3;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon, and Kidney;
RX      MEDLINE=22388257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA      Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA      Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Hsieh F.;
RA      Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA      Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA      Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA      Brownstein M.J.; Uedlin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA      Bosak S.A.; McKernan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA      Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Huiyk S.W.;
RA      Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA      Fahey J.; Helton E.; Kettelman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA      Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA      Blakeley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA      Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.W.; Butcherfield Y.S.;
RA      Krzywicki M.I.; Skalska U.; Smalins D.E.; Schnerch A.; Schein J.E.;
RA      Jones S.J.; Maizra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Kidney;
RA      Strausberg R.;
RN      Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon;
RA      Strausberg R.;
RN      Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon;
RA      Strausberg R.;
RN      Submitted (Jun-2002) to the EMBL/GenBank/DBJ databases.

```

RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon, and Kidney;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=liver;
 RX MEDLINE=22744052; PubMed=12859986; DOI=10.1016/S0003-9661(03)00286-8;
 RA Purihara T., Hosokawa M., Nakata F., Satoh T., China K.;
 RT "Purification, molecular cloning, and functional expression of
 RT inducible mouse liver acylcarnitine hydrolase in C57BL/6 mouse,
 RT belonging to the carboxylesterase multigene family.";
 RL Arch. Biochem. Biophys. 416:101-109(2003).
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL, BC015290; AAH15290.1; -;
 DR EMBL, BC024552; AAH24552.1; -;
 DR EMBL, BC031170; AAH31170.1; -;
 DR EMBL, BC034178; AAH34178.1; -;
 DR EMBL, BC034180; AAH34180.1; -;
 DR EMBL, BC034191; AAH34191.1; -;
 DR EMBL, AB110073; BAC76623.1; -;
 DR HSSP, F12337; IK4Y.
 DR MGD, MGI:2385905; Cee2.
 DR GO, GO:0005615; C:extracellular space; TAS.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE, PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE, PS00941; CARBOXYLESTERASE_B_2; 1.
 DR KMW, KMW:K00001; Signal.
 FT SIGNAL 1 26 Potential.
 FT CHAIN 27 561 carboxylesterase ML3.
 SQ SEQUENCE 561 AA; 62469 MW; F793967A646EB728 CRC64;
 Alignment Scores:
 Pred. No.: 1.08e-85 Length: 561
 Score: 1282.50 Matches: 276
 Percent Similarity: 61.93% Conservative: 77
 Best Local Similarity: 48.42% Mismatches: 174
 Query Match: 33.12% Indels: 43
 DB: 2 Gaps: 14
 US-10-023-515-1 (1-2158) x Q91W60 (1-561)
 QY 135 TTTTTCGATCTCTCCAGCCCTGTTGGGACACAGACAGTGGGAAATCTGGCCTTCT 194
 DB 15 PheGlyLeuLeuLeu-----LeuLeuLehIsvaGln--GlyGlnmserProGlu 31
 QY 195 GCTGAAGGCCACAGAGAACACAGGCTGGAGTGTTCAGGGGACAGCACTGCTG 254
 DB 32 Ala---AanProIleArgAsnThrHsThrGlyGlnIleGlnGlySerLeuLehIsva 50
 QY 255 CTGGAGAGCCCTGTCCTGTGAACGTCTTCTGGAGTCCCTTGTGCTGCTCCCGCTG 314
 DB 51 LysAspThrIlyAlaGlyAlaHsThrPheLeuGlyIleProPheAlaIlyAspProVal 70
 QY 315 GGATCCCTCGGATTAACGACCCGAGCCTGCATGCCCTTGGGATTACTTGGAGAAC 374
 DB 71 GlyProLeuArgPheAlaProProGluAlaProGluProTyrSerGlyValArgAspGly 90
 QY 375 ACCTCTCACTTAATTTGTGCTCCAGACTGAGTGGCTG-----CTCTTA 422
 DB 91 ThrAlaHsProAlaMeCysLeuGlnAenLeuAspMetLeuAsnGlnAlaGlyLeuPro 110
 QY 423 GATCAACACATCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGCTGCTTAC 482
 DB 111 AspMetLysMetMetLeuSerSerPhePro-----MetSerGluAspCysLeuTyr 127
 QY 483 CTGAACATCTATGCGCTGCGCCACAGCGGATACAGGTCTCAAGTCCCTCTTGTGTGG 542
 DB 128 LeuAsnIleTyrThrProAlaHsAlaHsGlyGlnSerAsnLeuProValMetValTyr 147
 QY 543 TTCACGAGAGTGCCTTCAAGACTGCGCTCAGCTCATCTTGTGATGGTCCGCGCTGCT 602

DB 148 IleHsGlyGlyAlaLeuValIleGlyMetAlaSerMetPheAspGlySerLeuLeuThr 167
 QY 603 GCCTATGAGACGCTGCTGTTGGTCCGATCAAGTCCGCTGAGATATTTGGTCTTC 662
 DB 168 ValAsnGluAspLeuValValThrIleGlnTyrArgLeuGlyValLeuGlyPhePhe 187
 QY 663 ACCACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACAGCTGCTGCTCT 722
 DB 188 SerThrGlyAspGlnHsIleAlaArgIlyAsnTyrGlyLeuAspGlnAlaIleAlaLeu 207
 QY 723 TCTGTGGTCCAGAAACATCAGTCTTCCGTGGGACCCAGCTCTGTACATCTTT 782
 DB 208 ArgTyrValGlnGlnAsnIleAlaHsPheGlyGlyAsnProAspArgValThrIlePhe 227
 QY 783 GCGAGTCCGGGGGAGCCATAGTGTCTTATCTGATCTGCTCCATGGCCAAAGGC 842
 DB 228 GlyGlnSerAlaGlyIlyThrSerValSerSerHsIsvaValSerProMetSerGlnGly 247
 QY 843 TTATTCACAAAGCCATCATGAGAGTGGGTGGCCATATCCCTTACCTGAGGCCCAT 902
 DB 248 LeuPheHsIsvaIleAlaMetCysSerGlyValAlaLeuLeuProAspLeuLeuSer--- 266
 QY 903 GATTATGAGAGAGTGGAGACCTGCAG--GTGTTGCACATTTCTGTGTAACTGCG 959
 DB 267 -----GluThrSerGluMetValSerThrThrValAlaIlyLeuSerGlyCysGluAla 284
 QY 960 TCAGACTGAGGCGCTGCTGAGTGGCTGAGGACAAACCTCCAGAGAGCTGCTGACC 1019
 DB 285 MetAspSerGlnAlaLeuValArgCysLeuArgGlyLysSerGluAlaGlyIleLeuAla 304
 QY 1020 CTCACCCAGAAACAAAGCTCTTCACTGAGTGTGTGATGGTCTTCTTCTTATGAG 1079
 DB 305 IleAsnIlyValPheIlyMetIleProAlaValAlaPheGlyGluPheProArgHs 324
 QY 1080 CCTTAGATCTATTTCTCAGAAACATTTAAAGCATTTCTTCATCATCTGAGTCAAT 1139
 DB 325 ProLysGluLeuLeuAlaSerGluAspPheHsIsvaProSerIleIleGlyAlaAsn 344
 QY 1140 AACCAAGAGTGTGGTCTGCTGCTGCTTGAAGAGAGGCTCCGATGCTCGTGGCTCC 1199
 DB 345 AsnAspGluPheGlyTyrSerIlePro-----ValValMetCylSer 358
 QY 1200 AAC-----AACTCCCTGCCCTCATGTATACAAACATCTGCAC----- 1241
 DB 359 AlaGlnMetIleGlyGlyIleThrArgGluAenLeuGlnAlaValLeuLysAspThrAla 378
 QY 1242 -----ATCCGCTCAGATTTGACCTTGTGGCTTAATGATATTCTTCATGAC 1289
 DB 379 ValGlnMetMetLeuProProGluCysSerAspLeuLeuMetGluGlyTyrMetGlyAsp 398
 QY 1290 AAGCACTCCCTGACCTGAATCCGAGACAGTCTTCTGACATTCCTTGGAGATGTGTTCTT 1349
 DB 399 ThrGluAspAlaGlnThrLeuGlnIleGlnPheThrGluMetMetCylAspPheMetPhe 418
 QY 1350 GTGTCTCCGCACTGATACAGCTGATATACACAGATGCTGGTGGACCTTCTTACTTC 1409
 DB 419 ValIleProAlaLeuGlnValAlaHsPheGlnHsG---SerHsAlaProValIlyThrPhe 437
 QY 1410 TATGAGTTCGGCACCGGCTCAGTGTCTTGAAGACAGAAACGAGCTTTGTGCAAGCC 1469
 DB 438 TyrGluPheGlnHsIleProProSerTyrPheLysAspValArgProProHsIvalIlyAla 457
 QY 1470 GACCAACGCTGATGAAGTCCGCTTGTGTGGTGGTGCCTTC-----CTGAAGGGGAC 1523
 DB 458 AspHsIsvaIleAspGluIleProPheValPheAlaSerPhePheTyrGlyMetLysLeuAsp 477
 QY 1524 ATTTGTTATGTTCAAGAGACCCAGGAGAGAGAGATTACTGAGCGGACGATGATGAAA 1583
 DB 478 Phe-----ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 491
 QY 1584 TACTGGGCTACTTGTCTGGAACCGGGAATCTTAATGGGAAACGACTGTCTCTGTGGCCA 1643
 DB 492 TyrTrpAlaAsnPheAlaArgHsIsvaIleGlyAsnProAsnSerGlnGlyLeuProTyrTrpPro 511

QY 1644 GCTTATATCTGACGTGAGGTAACCTCCAGCTGACTTGACATAGACCTCGGACAGAGA 1703
 Db 512 ValMetAspHisAlaSerGlnIleuGlnIleuAlaValGlyArgAla 531
 QY 1704 CTCAGAGAACCCGGGTGGATTTTGGACACGACCATCCCTCC-----CTGATCCTG 1754
 Db 532 LeuYsAlaGlyArgIleuGlnIleuPheThrIleuPheProGlnIleuIleuGlnIleu 551
 QY 1755 TCTGCTCCGACATGCTCCACAGTCTCTT 1784
 Db 552 LysAlaSerGlnAlaPheIleuArgIleu 561
 RESULT 10
 Q8BK48
 ID Q8BK48 PRELIMINARY; PRT; 559 AA.
 AC Q8BK48;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
 DE RIKEN full-length enriched library, clone:503145B19 product:stimilar
 DE to CARBOXYL-ESTERASE (EC 3.1.1.1) (AII-ESTERASE) (B-ESTERASE)
 DE (MONOUBUTYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
 DE (9030624102Rik protein).
 GN Name=9030624102Rik;
 OS Mus musculus (Mouse);
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kikunai T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Yoneda Y., Ishikawa T., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RA "RIKEN integrated sequence analysis (RISA) system-384-format

RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Ovary and uterus;
 RA Daichi U., Aizawa K., Akimura T., Arahata T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiroaka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata N., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akibira S., Takeda T., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altchul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Umedin T.B., Toshitoki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Guarnate P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Binkley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smalms D.E., Schermer A., Schen J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RA Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL; AK077248; BAC36707.1; -;
 DR EMBL; BC055062; AAH55062.1; -;
 DR HSSP; P12337; 1K4Y.
 DR MGD; MGI:2443170; 9030624102Rik.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR000379; Ser_CeSTR.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYL-ESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYL-ESTERASE_B_2; 1.
 KW Hydrolase.
 SO SEQUENCE 559 AA; 62317 MW; CCI4C61034A122C3 CRC64;
 Alignment Scores:
 Pred. No.: 1.4e-85 Length: 559
 Score: 1281.00 Matches: 260
 Percent Similarity: 63.48% Conservative: 79
 Best Local Similarity: 48.69% Mismatches: 165
 Query Match: 33.08% Indels: 30
 DB: Gaps: 8
 US-10-023-515-1 (1-2158) x Q8BK48 (1-559)
 QY 186 GGGCCCTTCTGAGGAGGACAGGACACGAGCTGGATGATTCAGGCGACAGCA 245
 Db 26 GlyGlnAspSerAlaSerProIleArgAsnThrHisThrGlyGlnValArgGlySerLeu 45

Alignment Scores:

Pred. No.: 3e-85 Length: 559
 Score: 1276.50 Matches: 262
 Percent Similarity: 64.58% Conserved: 79
 Best Local Similarity: 49.62% Mismatch: 172
 Query Match: 32.97% Indels: 15
 DB: 2 Gaps: 9

US-10-023-515-1 (1-2158) x 035533 (1-5559)

```

QY 177 GGAAGAAACCTGGGCTTTCTGCTGAGAGGCGCAGAGGAAACACAGGCTGGAGATTGAC 236
DB 26 GYGINAAPSerserProser-----ProIleargAenThrhIsthGlyGlnValArg 42
QY 237 GGCAGACAACTCATCTGTCTGGAGACCCCTGAGCTTGAACGCTTCTTCTGGAGTCCCG 296
DB 43 GYAGLeuValHLeValLysAAspThrhGlnValAspValHLeThrhPheLeuGlyIlePro 62
QY 297 TTTGCTGCTCCCGCTGGAGTCCCTGCGATTTCAGAACCCGACGCTGACGCTGG 356
DB 63 PheAlaLysProProValGlyProLeuAAspPheAlaProProGlnProProGlnProTrp 82
QY 357 GATACCTGGAGAGGACCACTCTACCTTAATTTGCTTCCAGAACTCAGAGTGGCTG 416
DB 83 SerGlyValAAspAAspGlyAAspSerPheProAlaMetCysLeuGlnAAspAAspMetMet 102
QY 417 CTCTTAAGT---CAACACATGCTCAGATGATTAACCGAAATTCCGAGTGTAGAAAGC 473
DB 103 AAspSerGlnGlyLeuLysMetIleLeuValIleMetProProIleSerMetSerGlnAAsp 122
QY 474 TGGCTTACCTGAAACATCTATGCGCTTCCGACGCGGATACAGAGCTTCCAGCTCCCGCT 533
DB 123 CysLeuLysLeuAAspHLeuLysTrpThrhHLeAlaHLeGlnGlySerAAspLeuProVal 142
QY 534 TTGGTGTGTCTCCAGAGAGTGCCTTCAGACATGCGCTCAGGCTCCATCTTTAGTGGCTCC 593
DB 143 MetValTrpIleHLeGlnGlyAlaLeuValGlyMetAlaSerMetCysAAspGlySer 162
QY 594 GCGCTGGCTGCTATAGAGACGCTGCTGTGTGCTGCTCAGATACCGGCTAGAAATATT 653
DB 163 MetLeuAlaHLeGlnAAspValIleValIleHLeGlnGlyAAspProGlnVal 182
QY 654 GGTTCCTTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713
DB 183 GlyPhePheSerThrhGlyAAspGlnHLeAlaArgGlyAAspTrpGlyTyrLeuAAspIleVal 202
QY 714 GCTGCTCTGCTCCGCTCAGAGAAACATGATTCCTTGTGAGGAGACCCGACTGTGTG 773
DB 203 AlaAlaLeuAAspTrpValGlnGlnAAspIleAlaHLePheGlyGlyAAspProGlnVal 222
QY 774 ACCATCTTGGGAGATCGCGGAGGACATAAGTTTCTAGCTTATCTGTCTGCTCCAGT 833
DB 223 ThrIlePheGlyGlnSerAlaGlyLysThrSerValIleSerHLeValIleSerProMet 242
QY 834 GCCAAAGGCTTATTCACAAAGCCATCATGAGATGGGAGTGGGAGCTCCCTTACCTG 893
DB 243 SerLysGlyLeuPheHLeGlyAlaIleMetGlnSerGlyValSerValMetProGlyIle 262
QY 894 GAGGCGCCATGATTATGAGAAAGAGTGAACATCGACGGTGTTCACATTTCTGTGTAA 953
DB 263 IleSerSerSerSerGlnMetVal-----TyrThrhIleValAlaAAspLeuSerAAspCys 280
QY 954 AATGGGTGAGACTCTGAGGCGCTGTGAGTGGCTGAGAGCAAAACCTCCAGAGACTG 1013
DB 281 AlaAlaValAAspThrhGlnThrhLeuValSerCysLeuAAspGlyLysSerGlnAlaGlnIle 300
QY 1014 CTGACCCCTCAGCCAGAAACAAAGCTTTTCACTGAGTGGTGTGAGTGGTCTTCTTCT 1073
DB 301 LeuAlaIleAAspLysValPheLysIleIleProAlaValIleAAspGlyGlnPheLeuPro 320
QY 1074 AATGAGCTCTAGATCTATTGTCTCAGAAAGCATTTTAAAGCAATCTCTTCATCATCGGA 1133
DB 1133

```

```

DB 321 LysHLeProGlnGlnLeuLeuAlaSerAlaAAspPheAAspProValProSerIleIleGly 340
QY 1134 GTCAATATACAGAGAGTGTGCTTCTGCTGCTCT-----ATGAGAGAGCTCTGAGATC 1187
DB 341 ValAAspAAspIleLysGlyLysPheLeuValProValIleMetGlySerAlaGlnIle 360
QY 1188 CTCAGTGGCTCCAGAACATGCTCTT-----GCCCTCATCTGATACAAACATCCTCGAC 1241
DB 361 LysGlnIleThrhArgValThrhLeuProAlaIleLeuLysSerThrhAlaAAspMetLys 380
QY 1242 ATCCGCGCTCAGATATTGACCTTGTGCTTAATGAATATCTTCATGACAGACATCCTCTG 1301
DB 381 LeuProProGlnCysGlyAAspLeuLeuMetGlnGlyLysMetGlyAAspAAspGlnAAsp 400
QY 1302 ACTGAATCCGAGACAGCTTCTGAGCTTGGCTGGAGATGTGCTTCTTGTGCTGCTTCA 1361
DB 401 GlnThrhLeuGlnAlaGlnPheArgGlnMetGlnMetGlyAAspPheThrhPheValIlePro 420
QY 1362 CTGATCACAAGCTCGATATATCAGAGAGATGCTGATGACCTGTCTACTTCTATGATTGCG 1421
DB 421 LeuGlnValAlaHLePheGlnArg---AlaHLeProValTyrPheTyrGlnPheGln 439
QY 1422 CACGCGCTCAGTGTGCTTGAAGACAGAGCGGCTTTGTCAAGCCGACACGCTGAT 1481
DB 440 HLeAAspProSerPhePheLysAAspThrhLysProHLeValIleAAspHLeGlyAAsp 459
QY 1482 GAACTCGCTTGTGTGTGCTGCTGCTCTTCTGAGGAGACATGTTATGTTATGAGAA 1541
DB 460 GlnPhePhePheValPheGlyAAsp---LeuLeuCysGly-----IleLysPhe---Ala 475
QY 1542 GCGACGAGAGAGAGAGAGTATCTGAGCGGAGAGATGATGAATATCTGAGGCTTGTGCT 1601
DB 476 SerThrhGlnGlnGlnGlnGlnLeuLeuSerAAspLysMetCysLysTrpAlaAAspHLeAla 495
QY 1602 CGAACCGGAAATCTTAATGGAACGACCTGTCTGTGCGGACGCTTATATATGACTGAG 1661
DB 496 ArgHLeGlyAAspProAAspSerGlnAAspLeuProTyrTrpProMetLeuAAspHLeAAsp 515
QY 1662 CAGTACCTTCAAGCTTGAACATGAGCTTGAACAGAGCTTGAACAGAGCTTGAACAGCGGCTG 1721
DB 516 GlnTyrLeuGlnLeuAAspIleLysProAlaValGlyArgAlaLeuLysValAAspArgLeu 535
QY 1722 GATTTTGGACAGACCATCCCC 1745
DB 536 GlnPheThrhThrhLysIleLeuPro 543

RESULT 12
QY 08BM97 PRELIMINARY; PRT; 559 AA.
AC 08BM97;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male colon cDNA, RIKEN full-length enriched
DE library, clone:9010624L02 product:similar to CARBOXYESTERASE (EC
DE 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
DE (PROCAINE ESTERASE) (METHYLEUTYRASE).
GN Name=9030624L02R1f;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA MEDLINE=99279253; PubMed=10343636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carinci P, Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX

```

RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RA The FANTOM Consortium,
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci F., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RT Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci F.,
 RT Kono H., Akiyama J., Nishi K., Kikunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci F.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka K., Hirozane T.,
 RA Hoti F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kachi H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AK033563; BAC28361.1; -.
 DR HSSP; P12337; 1K4Y.
 DR MGD; MGI:2443170; 9030624L02Rik.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR000379; Ser_ester.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 KW Hydrolase.
 SQ SEQUENCE 559 AA; 62305 MW; 8C54D21522C252C4 CRC64;

Alignment Scores:
 Pred. No.: 7.63e-85 Length: 559
 Score: 1271.00 Matches: 260
 Percent Similarity: 63.30% Conservative: 78
 Best Local Similarity: 48.69% Mismatches: 166
 Query Match: 32.83% Indels: 30
 DB: 2 Gaps: 8

US-10-023-515-1 (1-2158) x Q8BM97 (1-559)
 QY 186 GGGCTTCTGTGAAGGCGCACAGAGAAACCCAGCTGGATGTGATTCAAGGCAAGCA 245

DB 26 GYGLNAPSERSALASERPROLARGANTRH1STRNGYGLNVALARGLYSERLEU 45
 QY 246 GTCACTGTGCTGGGAAGCCCTGCTGAAAGCTTCTTGGAGTCCCTTGTGCT 305
 DB 46 VALNLSVALVYASPTASPLASPLALVALNLSRPHLEUGLYLEPROHEALALYS 65
 QY 306 CCCCCGCTGGGATCCCTGGATTTACGAACCCGAGCCGTGATCCGCTGGGATTA 365
 DB 66 PROVALGLYPROLEUDRPHENALPROFGLUALPROGLUPROTPROSERGLYVAI 85
 QY 366 CGAAGCCACCTCTTACCTTAATTTGGCTCCAGAACTCAGAGTGCTG--CTCTTA 422
 DB 86 ARGAPRGYLTHSRNHSRHSRHSRHSRHSRHSRHSRHSRHSRHSRHSRHSRHS 105
 QY 423 GATCAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGTCTTAC 482
 DB 106 ASPLAULYSMEVLEASPLAULYSMEVLEASPLAULYSMEVLEASPLAULYS 125
 QY 483 CTGAACATCTATGGCGCTCCGAGCGCATACAGGCTCCAGGCTCCGCTGGTGG 542
 DB 126 LEUASNLLEYRVALPROALAHVALHISGLUGLYSERASNLLEUVALHIS 145
 QY 543 TTCCGAGAGGTGCTTCAAGACTGCTCAGGCTCCATCTTGTGATGGTCCGCTG 602
 DB 146 ILEHISGLYGLYALALEUTHRVALGLYMETALASERMETRYASPGLYSERMET 165
 QY 603 GCTATGAGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
 DB 166 ALATNRGLUAPVALVALVALVALVALVALVALVALVALVALVALVALVALVAL 185
 QY 663 ACCAATGAGTACGACGCTCCGAGGAGTCCGCTTCAAGAGCAGGTGCTGCTG 722
 DB 186 SERTRGLYASPLAULYSMEVLEASPLAULYSMEVLEASPLAULYSMEVLE 205
 QY 723 TCTTGGTTCAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 782
 DB 206 ARGTRPVALGLNGLNLSRHSRHSRHSRHSRHSRHSRHSRHSRHSRHSRHS 225
 QY 783 GCGGAGTCCGCGGAGCCATTAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTT 842
 DB 226 GLYGUSERLALAGLYLTHSRVALSERHSRHSRHSRHSRHSRHSRHSRHSR 245
 QY 843 TTATTCACAAAGCATATGATGAGTGGGAGGATCCATCCCTTACTGAGGCCAT 902
 DB 246 LEUPHEHISGLYALALISMEVLEASPLAULYSMEVLEASPLAULYSMEVLE 265
 QY 903 GATTATGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 962
 DB 266 SERSERGLUMETVALHIS-----ARGILEVALALASNLLEUSERGLCYBALAL 283
 QY 963 GACTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 DB 284 ASNSERGLUTRLEUMETCYSEYSEUARGLYLSAANGLYLALGLUMETLEUAL 303
 QY 1023 AGCCGAGAAACAAAGTCTTCACTGAGTGTGATGATGATGATGATGATGAT 1082
 DB 304 ASNLVSVLPHLEYSLEILEPROGLYVALVALASPLAULYSMEVLEASPLA 323
 QY 1083 CTAGATCTATGCTCAGAGATTTAAAGCATTTTCAATCTTCTCATCATGAGTA 1142
 DB 324 GLNGULYSMEVLEASPLAULYSMEVLEASPLAULYSMEVLEASPLAULYS 343
 QY 1143 CACGAGTGGCTTCTGCTGCTGCT----- 1166
 DB 344 ASPLGLYRGLYTRPGLILEUPTHRLILEMETASPLAVALGLNLYLEGLULILE 363
 QY 1167 ---ATGAAGAGGCTCTGAGATCTCTAGTGGCTCCAAAGTCCCTGCTCATG 1223
 DB 364 THRARGLYSRHLEUPROVALVALLEU-----LYSETRHLEULYSMET 379
 QY 1224 ATACAAACATCTGACATCCGCTCAGTATTGTCACCTTGGTGAATGAATAC 1283


```

Db 380 Met-----LeuProProGluGlyGlyAspLeuLeuMetGluGluTyrMet 394
QY 1284 CATGACAGAGACTCCCTGACTGAAATCCGAGACAGCTCTTGACACTCTTGAGATGTG 1343
Db 395 GlyAspThrGluAspProGluThrLeuGlnIleGlnPheArgGluMetGlyAspPhe 414
QY 1344 TTTCTTGTGTCCTCGACAGTATGACAGCTGATATGACAGATGAGTGGTGGTGC 1403
Db 415 MetPheValIleProAlaLeuGlnValAlaIlePheGlnArg---SerHisAlaProVal 433
QY 1404 TACTTCTAGAGTTTGGGACACCGGCTTACAGCTTGTGAAGACAGAACCGGCTTTTGTG 1463
Db 434 TyrPheTyrGluPheGlnHisArgProSerPhePheLeuAspPheArgProTyrVal 453
QY 1464 AAGACCCGACAGCTGATGAGAGTCCGCTTGTGTCGTCGCTTCTCGAAGGGGAC 1523
Db 454 LysAlaIleAspHisGlyAspGluIlePheLeuValPheGluTyrGlnPhe-----GlyAsn 471
QY 1524 ATTGTATGTTTGGAGAGACCGACAGAGAGAGAGAGTACTGAGCCGGAAGATGATAA 1583
Db 472 IleLysLeu-----ProTyrThrGluGluGluGluGluGlnLeuSerArgArgMetMetLys 489
QY 1584 TACTGGGCTACCTTGTGTCGACCGGAGATCTTAATGGAAACAGCTGTCTGTGGCCA 1643
Db 490 TyrArgAlaAsnPheAlaArgHisGlyAsnProAsnSerGluGluLeuProTyrTyrPro 509
QY 1644 GCTTATATCTGACTGACAGTACTCTGACCTGAGTGAACATGAGCTCGACAGAGA 1703
Db 510 ValMetSerHisArgGlnGlnTyrLeuGlnLeuAspIleGlnProSerValGlyArgAla 529
QY 1704 CTCGAAGACCGCGGGGGATTTTGGACCGACAGACCATCCCC 1745
Db 530 LeuLysAlaArgArgLeuGlnPheTyrThrLeuPro 543

RESULT 13
ID 06PDB7 PRELIMINARY; PRT; 556 AA.
AC 06PDB7;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Hypothetical protein BC015286.
GN Name=BC015286;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.B.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., McKernan P.J., McKernan K.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Maitra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: BC058815; AAH58815.1; -.
DR HSSP: P12337; IK41.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser. esterase.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
DR KEGG: K00941; Carboxylesterase.
SQ SEQUENCE 556 AA; 61927 MW; 58DE05A3C73BECB CRC64;

Alignment Scores:
Pred. No.: 1,63e-84
Score: 1266.50
Percent Similarity: 60.17%
Best Local Similarity: 47.07%
Query Match: 32.71%
DB: 2 Gaps: 12

US-10-023-515-1 (1-2158) x 06PDB7 (1-556)
QY 96 ATGGCAGAGGAGCTTACTTCTGCTTCAAGATGCTTTTCTGATT-----CTC 149
Db 1 MetProArg-----SerGlnMetHisAsnThrLeuAspValIleLeuPheGluLeu 17
QY 150 CAGCCCTGTGTGGACACACAGTGGGAGAAAGTGGGCTTCTGTAAGGCGGACAG 209
Db 18 LeuLeuLeuLeuGlnHisValGln---GlyGlnAspSerProGluAla---SerProIle 35
QY 210 AGGAACACAGAGCTGGGATGATTCAGGGGACAGCAAGTCACTGTGCGAAGCCCTGTG 269
Db 36 ArgAsnThrHisThrArgGlnValArgGlySerLeuValHisValLysAsnThrLysVal 55
QY 270 CCTGTGAAGCTGTTCCTCGGAGTCCCTTCTGCTGCTCCCGCTGGAGATCCCTGCAATT 329
Db 56 GlyAlaHisThrPheLeuGlnIleProPheAlaLysProProValGlyProLeuArgPhe 75
QY 330 ACGAACCCGACGCTCATGCGCCCTGGAGTAACTTGCAGAAAGCAAGCTCTACCCCTAT 389
Db 76 AlaProProGluIleAspProGluProTyrPheSerGlyValArgAspGlyThrAlaHisProAla 95
QY 390 TTGGGCTCCAGAACTCAAGTGGCTCTTATGATCAACATGCTCAAGGTGCAATTAC 449
Db 96 MetCysLeuGlnAsn-----LeuGlyValMetLysGluIleLysLeuLeu 111
QY 450 CCGAAATTCGAGTGTCAAGAGATGCTCTTACCTGAGACATGATGAGCTGGCCGACGCC 509
Db 112 ProProValSerThrSerGluAspGlyLeuIleValIleTyrThrProAlaHisAla 131
QY 510 GATACAGGCTCCAGAGTCCCGCTTGTGTGTGTTCCAGAGAGTCCCTTCAAGAGTGGC 569
Db 132 HisGluGlySerAsnLeuProValMetValTrpIleHisGlyGlyLeuValAlaLeu 151
QY 570 TCAGCTCCATCTTGAATGAGTGGTGGCTGGCTGCTTGAAGAGCTGCTGTGTGTC 629
Db 152 MetAlaSerMetLysAspGlySerLeuLeuAlaAlaIleGluAspLeuValValThr 171
QY 630 GTCCAGTACCGGCTTGAATATTTGGTTTCTTCAACATGAGGATGACATGCTCCGGG 689
Db 172 IleGlnTyrArgLeuGlyValIleGluIlePheSerThrArgLysArgGlnHisAlaArgGly 191
QY 690 AACTGGGCTTCAAGAGACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
Db 192 AsnThrGlyPheLeuAspGlnValAlaAlaLeuLysArgTrpIleGlnGlnAsnIleAlaHis 211
QY 750 TTGGTGGGAGCCCGAGCTCTGTGACCATCTTGTGGAGTCCGCGGAGCATTAAGTTT 809
Db 212 PheGlyGlyLysProAspArgValThrIlePheGluGluSerAlaGlyGlyThrSerVal 231

```



```

QY 810 TCTAGCTTACTGTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAT 869
DB ||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 232 SerSerHisValValSerProMetSerLySGlyLeuPheHisGlyAlaIleMetGlySer 251
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 870 GGGGTGGCCATCTTACCTGAGAGCCATGATTAAGAAAGTAGAGAGAGAGAGAG 929
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 252 GlyValAlaLeuLeuProTyrLeuIleThrAspThrSerGluMetValSer-----Thr 269
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 930 GTGGTTGCATTTCTGTGTGAACATGGCTGAGACTCTGAGAGCCCTGCTGAGAGTCCG 989
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 270 ThrValAlaValLeuSerGlyCysGlyAlaMetAspSerGlyAlaLeuValArgCysLeu 289
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 990 AGGAGAAAACCTCCAGAGAGGCTGACCTCAGCAGAGAAAACAAAGCTTTCATCTGCA 1049
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 290 ArgGlyLySerGlyAlaGlyIleLeuAlaIleValLeuValGlyMetIleProAla 309
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1050 GTGGTTGATGTGCTTTCTTCTTAATGAGCCTTAAGATCTATTTGCTGAGAAAGCATTT 1109
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 310 ValValAspGlyGlyPhePheProArgHisProLyGlyLeuLeuAlaSerGlyAspPhe 329
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1110 AAAGCAATCTTCCATCATGAGAGCATTAACAGAGAGTGGCTTCTGCTGCTTAAAG 1169
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 330 HisProValProSerIleIleGlyValAlaMetAspGlyPheGlyTyrThrIlePro--- 348
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1170 AAGAGAGCTCTGAGATCTCAGTGGCTCCAAAGAGTCCCTGAGCTCATGATCAAA 1229
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 349 -----ValValMetGlySerAlaGlnThrIleLeuGlyIleThrArgGly 363
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1230 AACATC-----CTGACATCCCGGCTCAGATATTTG 1259
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 364 AsnLeuGlnAlaValLeuLyAsnThrThrAlaGlnLeuMetLeuProGlyCysSer 383
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1260 CACCTTGGGTAATGAATCTTCCATGACAAAGCACTCCCTGACATGAAATCCGAGCAAT 1319
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 384 AspLeuLeuMetGlyGlyIleTyrMetGlyAspThrGlyAspAlaGlnThrLeuGlnIleGln 403
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1320 CTTTGGAGTCTTGGAGATGAGTCTTGTGGTCCCTGAGCCTGATCAAGCTGATAT 1379
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 404 PheThrGluMetMetGlyAspPheMetPheValIleProAlaLeuGlnValAlaTyrPhe 423
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1380 CACAGAGATGTGTGAGCTGTCTACTTCTATGATTTGCGACCGGCTCAGTGGCTTT 1439
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 424 GlnArg-----SerHisAlaSerValTyrPheTyrGlnPheGlnHisGlnIleAlaSerLeu 442
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1440 GAAAGACAGAGACCGGCTTTTGTCAAGCCGACAGCTGATGAGTCCGCTTGTGCTTC 1499
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 443 LyAspAspValArgProThrHisValLyAlaAspHisAlaAspGlyIleProPheValPhe 462
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1500 GGTGGTGCCTTC-----CTGAAGGGGGAATTGTTATTTGGAAGAGACCGAGAGAG 1553
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 463 GlyTyrPhePheThrAspMetLyLeuAspPhe-----ThrGlyGly 476
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1554 GAGAGATTACTAGAGCCGGAAGATGATGAATATCTGGGCTACCTTGTGCTGAAACCGGAGAT 1613
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 477 GlyLybLeuLeuSerArgHisGlyMetCysTyrTyrPheAlaAspPheHisGlyAlaAsn 496
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1614 CCTAATGGAGAACACTGTCTCTGTGGCCAGCTTATTAATCTGAATGAGAGTACCTCCAG 1673
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 497 ProAsnSerGlyGlyLeuProTyrTyrProValIleMetAspHisAspGlyGlnTyrLeuGln 516
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1674 CTGAGCTTGAACATGAGCTCGGAGCAGAGATCAAAAGAACCGGCGGTGATTTTGGAGC 1733
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 517 LeuAspThrGlnProAlaValGlyArgAlaLeuLySerArgTyrGlyLeuGlnPheTyrPhe 536
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1734 AGCACCATCCCG-----CTGATCTGTGTGCTGCTCGACATGCTCCACAGTCCCTTT 1784
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 537 LybThrLeuSerGlybIleGlnGlybArgAlaSerGlybAspLybHisThrGlybLeu 556
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

DT 05-JUL-2004 (TRENBLrel. 27, last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, last annotation update)
DE LOC443703 protein (Fragment).
GN Name=LOC443703;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Collins R.S., Wagner L., Shenmen C.M., Schler G.D.,
RA Klausner R.D., Collins P.S., Wagner L., Scherf C.F., Bhat N.K.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locantello N.A., Peters G.J., Abramson R.D., Mollaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smalios D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gernard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
DR EMBL: BC074230; AAH74230.1; -.
DR GO: GO:0016787; P:hydroxylase activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR003379; Ser_esterase.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydroxylase.
FT NON_TER 1 1
SQ SEQUENCE 568 AA; 63032 MW; 9336DDB6B53931A1 CRC64;

```

Alignment Scores:

```

Pred. No.: 2,74e-84 Length: 568
Score: 1263.50 Matches: 258
Percent Similarity: 64.45% Conservative: 87
Best Local Similarity: 48.22% Mismatches: 169
Query Match: 32.63% Indels: 21
DB: Gaps: 8

```

US-10-023-515-1 (1-2158) x Q6GMS4 (1-568)

```

QY 177 GGAAGAACTGGGCTTGTGCTGAAGGGCCACAGAGAACACCAAGCTGGATGATTCAG 236
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 28 GlyIleTyrGlyThrGlnAspAlaArgProLeuLeuThrThrAsnTyrGlyGlnLeuLeu 47

```



```

QY 96 ATGCCAGAGGA---CTTACTTCACTGCTTCAACATGGTCTTTTCTGATCTCCAG 152
DB 1 MetProhArganGlnLeuHisSerTrpLeuAsnAlaValLeuPheGlyLeuLeuLeu--- 19
QY 153 CCCCTGTTGGGACACAGACAGTGGGGAAMAACCTGGGCTTCTGTGAAGGCCACAGAG 212
DB 20 ---LeuLeuIleHisValGln---GlyGlnAspSerProGlnSer---SerTrpIleArg 36
QY 213 AACACACAGGCTGGGATGTGATTCAGGACAGACAGTCTGTCTGGGAAGCCCTTGCT 272
DB 37 ThrHisIleThrGlnGlyGlnValGlnGlyLeuAspRhiSValIlyAspRhiThrySAlaGly 56
QY 273 GTGAACGCTGCTCTGGAGTCCCTTTGGTCTCCCGGCTGGAGTCCCTGAGATTTAG 332
DB 57 ValHisThrPheLeuGlyIleProPheAlaLysProProValGlyProLeuArgPheAla 76
QY 333 AACCCGACGCTGCATCCGCTGGAGTAACTTGCAGAGACCACTCCCTACCTAATTTG 392
DB 77 ProProGlnProProGlnProTrpSerGlyValArgAspRhiAlaThrSerGlnProAlaMet 96
QY 393 TGCTTCCAGAACTCAGAGTGGCTG-----CTTTAGATCAACATGCTCAAG 440
DB 97 CysLeuGlnAsnLeuAspRilLeuLeuArgGlyValGlyLeuLeuAspMetLysMetIle--- 115
QY 441 GTGCATTACCCGAATTCGAGAGTGCAGAGACCTGCTTACCTGACATCTAAGCTGCT 500
DB 116 -----LeuSerSerIleSerMetSerGlnAspCysLeuTyLeuAsnAlaTyxAlaPro 133
QY 501 GCCACGCGGATACAGAGCTCCAAAGCTCCCGCTTGTGTGGTCCACAGAGTCCCTTC 560
DB 134 AlaHisAlaArgGlnGlySerHisLeuProValMetValTrpIleHisGlyIlyAlaLeu 153
QY 561 AAGACTGCTCAGCTCCATCTTTATGAGTCCGCTCCGCTGGCTGTATGAGAGCTGCTG 620
DB 154 ValValGlyMetAlaSerMetTyxAspGlySerLeuLeuThrValAsnGlyAspLeuVal 173
QY 621 GTTGTCGCTGCTCCAGTACCGGTAGGAATATTTGGTTTCTTCAACAATGGATACAGAT 680
DB 174 ValValThrIleGlnTyxArgLeuGlyValLeuGlyPhePheSerThrGlyAspGlnHis 193
QY 681 GCTCCGAGAACTGGGCTTCAAGGACAGGAGTGGCTGTCTGCTGGGTCAGAGAAC 740
DB 194 AlaArgGlyAsnTrpGlyTyxLeuAspRhiValAlaAlaLeuAspTrpValGlnGlnAsn 213
QY 741 ATCGAGTCTTGGTGGGAGACCCAGCTGTGACATCTTTGGGAGTCCGCGGAGCC 800
DB 214 IleAlaHisPheGlyGlyAsnProAsnArgValThrIlePheGlyGlnSerAlaGlyIly 233
QY 801 ATAAGTCTTCTAGCTTAATACGTCTCCATGCGCAAGGCTTATTCACAAAGCCATC 860
DB 234 ThrSerValSerSerHisValIleSerProMetSerGlnGlyLeuPheHisGlyAlaIle 253
QY 861 ATGAGAGTGGGAGGAGCCATCCCTTACCTGAGAGCCCATGATATGAGAAAGTGA 920
DB 254 MetGlnSerGlyValAlaLeuLeuProAspRleIleSer-----GlnThrSerGln 270
QY 921 GACSTGAG---GTGGTGCACATTTCTGTGGAACATGCTGACAGTCTGAGGCCCTG 977
DB 271 ThrValSerThrThrValAlaLysLeuSerGlyCysGlyAlaMetAspSerGlnAlaLeu 290
QY 978 CTGAGAGTCTGAGACAAACCCCTCCAAAGAGCTGACCTGACGACCAAGAAACAG 1037
DB 291 ValArgCysLeuAspAlaLysSerGlyAlaGlnIleLeuValIleAsnLysValaPheLys 310
QY 1038 TCTTCACTGAGAGTGTGATGCTGCTTCTTCTTAAGAGCTCTAGATCTATGCT 1097
DB 311 MetIleProAlaValaValaAspGlyGlyPheLeuProAlaArgHisProLysGlyLeuLeuAla 330
QY 1098 CAGAAAGCATTTAAAGCAATTTCTTCCATGATGAGAGTCAATACCAACAGATGAGCTTC 1157
DB 331 SerGlnAspRhiAspProAlaProSerIleIleGlyValaAsnThrAspRilutyTyxAsp 350

```

```

QY 1158 CTGCTGCTATG-----AAGAGGCTCCTGAGATCTCACT 1193
DB 351 ThrIleProMetValMetGlyTyxAlaGlnIleLeuLysGln-----LeuSer 366
QY 1194 GAGTCCAAACAGTCCCTTGGCTCCATGTATACAAACATCTGACATCCCGCTCAG 1253
DB 367 ArgGlnAsnLeuGlnAlaValLeuLysAspThrAlaAlaGlnMetMetLeuProGln 386
QY 1254 TATTTGACACCTTGGCTAATGATATCTTCATGACAGACATCCCTGATGAATCCGA 1313
DB 387 CysGlyAspRleLeuMetGlnGlyIlyMetGlyAsnThrAspAspSerGlnThrLeuGln 406
QY 1314 GACAGCTTCTGACATCTGCTGAGAGTGTCTTGTGTGCTCCCTGACATGATCAAGCT 1373
DB 407 IleGlnTyxThrGlnMetMetGlyAspRheLeuPheValIleProAlaLeuGlnAla 426
QY 1374 CGATATCAACAGAGTGTGTCACCTGTCTATCTTATGAGTTTGGACCGGCTCAG 1433
DB 427 HisPheGlnArg---SerHisAlaProValTyxPheTyxGlyPheGlnHisAlaProSer 445
QY 1434 TGCTTTGAAGACAGAACCCGCTTTGTCAAGCCGACCAACGCTGATGAAGTCCGCTT 1493
DB 446 TyxPheLysAsnValArgProProHisValLysAlaAspRhiSAlaAspGlyValaProPhe 465
QY 1494 GTGTGCGGTGCTCTTCTGTAAGGGGAGACATGTTATGTTGTAAGAGCCACGAGAG 1553
DB 466 ValPhe---GlySerPhePheSerGlyMetLysLeuAspPhe-----ThrGlnGly 481
QY 1554 GAGAAGTACTGAGCCGGAAGATGATGAATATCTGAGCTTACTTGTGTAACCGGAGAT 1613
DB 482 GlnArgLeuLeuSerArgArgMetMetLysTyxTrpAlaAsnPheAlaArgGlnGlyAsn 501
QY 1614 CCTATGGAAGACACTGTCTGTGCGCACCTTATCTGACCTGACGACATCACTCCAG 1673
DB 502 ProAsnSerGlnGlyLeuProTyxTrpProAlaLeuAspRhiSAspGlnGlnTyxLeuGln 521
QY 1674 CTGAGCTTGAACATGAGCTCGGACAGAGACTCAAGAGACCGGAGTGGATTTTGAAC 1733
DB 522 LeuAspThrHisProAlaValaAspArgAlaLeuLysValaArgLysLeuGlnPheThrThr 541
QY 1734 AGCACCATCCCC 1745
DB 542 LysThrLeuPro 545

```

Search completed: June 15, 2005, 08:52:45
 Job time : 226.639 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 07:48:14 ; Search time 32.6132 Seconds

(without alignments)
9878.981 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 3872

Sequence: 1 ccaacgcctccgaataaacagc.....aaaaaaaaaaaaaaaaaaaaa 2158

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Fgapop 10.0 , Fgapext 0.5
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_n2p.model DEV=xlh
-Q=/cgn2_1/USPTO_spool/6664091/runat_14062005_133500_9360/app_query.fasta.1.4238
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091@cgn2_1.1.54@runat_14062005_133500_9360 -NCPU=6 -ICU=3 -NO_MMAP
-LARGEQUERY -NEG_SCORES=0 -WAIT -DESBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*\n2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*\n3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*\n4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*\n5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*\n6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3079	79.5	581	4	US-10-023-515-2
2	1230	31.8	559	4	US-09-595-682B-28
3	1230	31.8	559	4	US-09-949-016-6426
4	1230	31.8	577	4	US-09-949-016-9670
5	1150.5	29.7	565	4	US-09-595-682B-21
6	1138.5	29.4	566	3	US-09-264-737-2
7	1130.5	29.2	543	4	US-09-595-682B-26
8	1091.5	28.2	584	2	US-08-845-285A-2
9	1091.5	28.2	584	3	US-09-140-933-2
10	1091.5	28.2	584	3	US-09-146-661-2
11	1091.5	28.2	584	3	US-09-150-515-2
12	1034.5	26.7	539	3	US-09-264-737-1

13	942.5	24.3	454	3	US-08-446-100-26	Sequence 26, Appl
14	942.5	24.3	454	3	US-08-446-100-28	Sequence 28, Appl
15	942.5	24.3	454	3	US-08-446-100-30	Sequence 30, Appl
16	942.5	24.3	454	3	US-08-446-100-31	Sequence 31, Appl
17	940.5	24.3	454	3	US-08-446-100-27	Sequence 27, Appl
18	940.5	24.3	454	3	US-08-446-100-29	Sequence 29, Appl
19	871	22.5	574	4	US-10-023-515-4	Sequence 4, Appl1
20	800.5	20.7	933	4	US-09-949-016-8386	Sequence 8386, Ap
21	789.5	20.4	933	4	US-09-949-016-8387	Sequence 8387, Ap
22	786	20.3	853	4	US-09-949-016-6888	Sequence 6888, Ap
23	777.5	20.1	843	4	US-09-491-356C-20	Sequence 20, Appl
24	777	20.1	575	1	US-08-348-920-1	Sequence 1, Appl1
25	775.5	20.0	823	4	US-09-491-356C-23	Sequence 23, Appl1
26	774	20.0	575	1	US-08-348-920-2	Sequence 2, Appl1
27	772	19.9	848	4	US-09-491-356C-22	Sequence 22, Appl
28	769	19.9	836	4	US-09-491-356C-21	Sequence 21, Appl
29	765.5	19.8	614	3	US-08-446-100-25	Sequence 25, Appl
30	764.5	19.7	614	1	US-07-732-962A-2	Sequence 2, Appl1
31	764.5	19.7	614	3	US-08-370-156-2	Sequence 2, Appl1
32	764.5	19.7	614	3	US-08-446-100-19	Sequence 19, Appl
33	764.5	19.7	614	3	US-08-446-100-21	Sequence 21, Appl
34	764.5	19.7	614	3	US-08-814-095-2	Sequence 2, Appl1
35	764.5	19.7	614	5	PCT-US92-06106-2	Sequence 2, Appl1
36	764.5	19.7	645	4	US-09-949-016-7063	Sequence 7063, Ap
37	764.5	19.7	645	4	US-09-949-016-7064	Sequence 7064, Ap
38	762.5	19.7	614	3	US-08-446-100-20	Sequence 20, Appl
39	761.5	19.7	614	3	US-08-446-100-23	Sequence 23, Appl
40	760.5	19.6	614	3	US-08-446-100-22	Sequence 22, Appl
41	759.5	19.6	600	2	US-08-370-156-4	Sequence 4, Appl1
42	759.5	19.6	600	3	US-08-814-095-4	Sequence 4, Appl1
43	759.5	19.6	600	3	US-08-975-084-1	Sequence 1, Appl1
44	755.5	19.5	617	2	US-08-370-156-6	Sequence 6, Appl1
45	755.5	19.5	617	3	US-08-814-095-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Cytis, Rory A. J.
; TITLE OF INVENTION: 5310, A NOVEL HYDROLYTIC ENZYME
; TITLE OF INVENTION: 5310, A NOVEL HYDROLYTIC ENZYME
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2
Alignment Scores:
Pred. No.: 2.11e-292
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 79.52%
DB: 4
Caps: 0
US-10-023-515-1 (1-2158) x US-10-023-515-2 (1-581)
QY 96 ATGCACAGGACATCTTCTGCTTCAACATGCTTTTCTGATTCACAGCC 155
|||||

```

Db      1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCy8PhePheLeuIleLeuGlnPro 20
QY      156 CTGTTGGGACACAGACAGTGGGGAAAACTGGGCTTTCTGTGAAGGGCCACAGAGAAC 215
Db      21 LeuLeuGlnYHSArGlnTrpGlySerThrGlyProSerAlaGlnGlyProGlnArGAsn 40
QY      216 ACCAGGCTGGATGATTCAAGGCAAGCAAGTCACTGTGGCTGGAGGCCCTGGCTGTG 275
Db      41 ThrArGLeuGlnTrpIleGlnGlySerGlnValThrValLeuGlnSerProValProVal 60
QY      276 AACGCTTCCTGGAGTCCCTTGTGCTGCTCCCGCTGGAGTCCCTGCATTTACGAAC 335
Db      61 AsnValPheLeuGlnYValProPheAlaIaIaProProLeuGlnSerLeuArGpHeThrAsn 80
QY      336 CCGAGACCTGCATGAGCCCTGGGATTAATCTGGAGAAACCACTCTTAACCTTAATTTGTC 395
Db      81 ProGlnProAlaSerProTrpPAspAsnLeuArGlnAlaThrSerTrpProAsnLeuCy8 100
QY      396 CTCCAGAACTCAGATGGCTGCTCTTAGATCAACACATGCTCAAGGTGATTAACCGAAA 455
Db      101 LeuGlnAsnSerGlnTrPLeuLeuLeuAspGlnHISmetLeuYsValHISTrpGln 120
QY      456 TTGCGAGTGCAGAAAGCTGCTTACTGAACATCTATGCGCTGCCACGCCGATACA 515
Db      121 PheGlyValSerGlnAspCy8LeuTrpLeuAsnHISLeTrAlaProAlaHISAlaAspThr 140
QY      516 GGCTCCAGACTCCCGCTTGTGTGTGTGTCCAGAGGTGCTTCAAGAAGTGGCTCAGGC 575
Db      141 GlySerYsLeuProValIleuValTrpPheProGlyGlyAlaPheYsThrGlySerAla 160
QY      576 TCATCTTTGATGGTCCGCGCTGGCTGCTGCTTATGAGAGACGTGCTGTGTGCTCCAG 635
Db      161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValGln 180
QY      636 TACCGCTAGAAATATTTGGTTCTTCAACACATGGAGTACAGATCCCGGAGAACG 695
Db      181 TrpArGLeuGlyIlePheGlnPhePheTrpTrpPAspGlnHISAlaProGlyAsnTrp 200
QY      696 GCCTTAAGACCAAGGTGCTGCTGTCTGCTGGGTCAGAGAACATCGAGTTCTTCGAT 755
Db      201 AlaPheYsAspGlnValAlaIaIaLeuSerTrpValGlnYsAsnHISGlnPheGln 220
QY      756 GGGAGCCCGAGCTGTGACCATCTTTGGCGAGTCCGCGGAGCCATAAGTCTTCTAGT 815
Db      221 GlyAspProSerSerValThrIlePheGlnGlnSerAlaGlyAlaIleSerValSerSer 240
QY      816 CTTAATCTGTCTCCATGCGCAAGGCTTAATTCACAAAGCATCATGAGAGTGGAGTG 875
Db      241 LeuIleLeuSerProMetAlaYsGlyLeuPheHISYsValIleMetGlnSerGlyVal 260
QY      876 GGCATCATCCCTTACTGAGGCGCCATGATTATGAGAAAGTGAAGACTCGAGGTGTT 935
Db      261 AlaIleIleProTrpYsLeuGlnAlaHISAspTrpGlnYsSerGlnAspLeuGlnVal 280
QY      936 GCACATTTCTGTGTAAACATGCGCTCAGACTCTGAGCGCTGAGAGTGTCTGAGACA 995
Db      281 AlaHISpHeCy8GlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArGlyCy8LeuArGTrp 300
QY      996 AAACCTTCAGAGGAGCTGCTGACCTCAGCCAGAAAACAAAGTCTTTCATCGAGTGT 1055
Db      301 YsProSerYsGlnLeuLeuThrLeuSerGlnYsThrYsSerPheTrpArGValVal 320
QY      1056 GATGGTGTCTTCTTCTTAATGAGCCTTAGATCTATGCTCTCAAGAAAGATTTAAAGCA 1115
Db      321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnYsAlaPheYsVal 340
QY      1116 ATTCCTTCATCATCGAGATCAATTAACACAGAGTGTGGCTTCTGCTGCTCATGAAGAG 1175
Db      341 IleProSerIleIleGlyValAsnAsnHISGlnYsGlyPheLeuLeuProMetYsGln 360
QY      1176 GCTCTGAGATCTCAGTGGCTCAACAGATCCCTTGGCTCATCTGATCAAAACATC 1235
Db      361 AlaProIuIleLeuSerGlySerAsnYsSerLeuAlaLeuHISLeuIleGlnAsnIle 380

```

```

QY      1236 CTGACATCCCGCTCAGATTTTGCACCTTTGGCTTAATGAATCTTCCATGACAGAC 1295
Db      381 LeuHISIleProProGlnTrpLeuHISLeuValAlaAsnGlnYsPheHISAspYsHIS 400
QY      1296 TCCCTGATGAATCCGAGACAGTCTTGTGACTTGTGGAGATGATGTTCTTTGTGCTC 1355
Db      401 SerLeuThrGlnIleArGAspSerLeuLeuAspLeuGlnYsPheValPhePheValVal 420
QY      1356 CCTGACATGATACAGCTCGATATACAGAGATGCTGTGACACCTGTACTTATGAG 1415
Db      421 ProAlaLeuIleThrAlaArGTrpHISArGAspAlaGlnAlaProValTrpPheTrpGln 440
QY      1416 TTTCGGCACCGGCTCAGTGTGTTGAAGACAGAAAGCGGCTTTTGTCAAAGCCGACAC 1475
Db      441 PheArGHIleArGProGlnCy8PheGlnAspTrpYsProAlaPheValYsAlaAspHIS 460
QY      1476 GCTGATGAATGCCGCTTGTGTGTGCTGCTGCTTCTGAAAGGGGACATTTGTTATG 1535
Db      461 AlaAspGlnValArGpPheValPheGlyGlyAlaPheLeuYsGlyAspIleValMetPhe 480
QY      1536 GAAGAGCCACAGAGAGAGAGAGATTAATGAGCCGAGAACATGATGAATCTGGGCTACC 1595
Db      481 GlnGlyAlaThrGlnGlnGlnYsLeuLeuSerArGlySmetMetYsTrpTrpAlaThr 500
QY      1596 TTGCTGAAACCGGGAATCCTTAATGGAGACACCTGTCTGTGGCCAGCTTAATCTG 1655
Db      501 PheAlaArGThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
QY      1656 ACTGAGAGATACCTCCAGCTGTGATGAACATGAGCTCGAGACAGACTCAAGAACCG 1715
Db      521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArGLeuYsGlnPro 540
QY      1716 CGGCTGATTTTGAACACAGACCAATCCCTGTGATCTGTGCTGCGCCGACATGCTCCAC 1775
Db      541 ArGValAspPheTrpHISerThrIleProLeuIleLeuSerHISAspMetLeuHIS 560
QY      1776 AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCTCCTCAGCCTTTCTTTCTTTGTGCT 1835
Db      561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCy8Ala 580
QY      1836 CCT 1838
Db      581 Pro 581

RESULT 2
US-09-595-682B-28
; Sequence 28, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR FILING DATE: 1999-02-12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 559
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-595-682B-28

Alignment Scores:
Pred. No.: 2,76-111 Length: 559
Score: 1230.00 Matches: 248

```

Percent Similarity:	63.84%	Conservative:	91
Best Local Similarity:	46.70%	Mismatches:	168
Query Match:	31.77%	Indels:	24
DB:	4	Gaps:	8

```

DB: 4 Gaps: 8
US-10-023-515-1 (1-2158) x US-09-949-016-6426 (1-559)
QY 186 GGGCTTCTGCTGAAAGGCCACAGAGAAACACAGAGTGGATGATTCAGAGGCAAGCA 245
DB 26 GYGINAASPserLaserProileArgThrThrHnIstHnGInValLeuGlySerleu 45
QY 246 GTCACTGTGTGCTGGAGCCCTGTGCTGAAACGTTTCTCGAGAGTCCCTTGTGCT 305
DB 46 ValHnSValLySglYAlaHnAlaGlyAlGInThrPheLeuGlyIleProPheAlaLyS 65
QY 306 CCCCCGTGGAGATCCCTGCGATTACAGAACCCGAGCTGCATGCGCCCTGGATACGTG 365
DB 66 ProPheLeuGlyProLeuArgPheHnIarProProGInuProProGInuSerTyrSerGlyVal 85
QY 366 CGAAGAGCCACCTCTACCTTAATTTGTGCTCCAGAAC-----TCAGAG 410
DB 86 ArgHnArgLyHnThrHnSProAlaMetCyLeuGInAASPLeuThrAlaValGInuSerGly 105
QY 411 TGAGTGTCTTGAATCAACATGCTCAAGGTGATACCGGAATTCGAGAGTCTAGAA 470
DB 106 PheLeu-----SerGInPheAhnMetThrPheProSerAASPserMetSerGly 121
QY 471 GAGTCCCTTCTTCAACATGATATGCGCTGCGCCACGCGATACAGGCTCCAGAGTCCGC 530
DB 122 AAsPyAAsLeuYrLeuSerIleYrThrProAlaHnSerHnSglInuGlySerAAsLeuPro 141
QY 531 GTCTTGTGTGTGCTCCAGAGAGGCTTCAAGAGTGGCTCAAGCTCCATCTTGAATGG 590
DB 142 ValMetValItrIleHnSglYAlaLeuValPheGlyMetAlaSerLeuYrAAsGly 161
QY 591 TCGGCTGCTGCTGCTTGAAGAGCTGCTGTTGTGCTGCTCCAGTACCGGCTAGAGATA 650
DB 162 SerMetLeuAlaIleuGInuAAsnValIleValIleIleGInuYrArgLeuGlyVal 181
QY 651 TTGTGTTTCTTCAACATGAGATGAGAGCTGCTGCTGCTGCTCCAGTACCGGCTAGAG 710
DB 182 LeuGlyPheHnSerThrGlyAAsPyHnSglYAlaThrGlyAAsnItrIleuAAsPIn 201
QY 711 GTGGCTGCTGCTGCTGCTGCTCCAGAGAAACATGAGATTTCTGAGGAGACCCAGCTCT 770
DB 202 ValAlaIleAAsnArgTrValGInuGInAAsnIleAlaHnIlePheGlyYrAAsnProAAsArg 221
QY 771 GTGACCATCATCTTGGAGAGTCCGGGAGGACCATAGTGTTCATGCTTAACTGTCTCC 830
DB 222 ValThrIlePheGlyGInuSerAlaGlyGlyHnSerValSerSerLeuValIleSerPro 241
QY 831 ATGGCCAAAGCTTATTCACAAAGCATATGAGAGAGTGGGCTGACCATCCCTTAC 890
DB 242 ILeSerGInuGlyLeuPheHnSglYAlaIleMetGInuSerGlyValAlaLeuLeuProGly 261
QY 891 CTGAGAGGCCCATGATTAGAGAAAGTGAAGACCTG-----CAGGTGTTCACATTTTC 944
DB 262 LeuIleAla-----SerSerIleAAsPValIleSerThrValIleAlaIleAAsnLeu 277
QY 945 TGTGTGAACAATGCTCAGACTCTGAGGCTGCTGAGAGTGGCTGAGAGCAAAACCTCC 1004
DB 278 SerAlaCyAAsPInValAAsPserGInuAlaLeuValGlyCySLeuAArgGlyYrSerLyS 297
QY 1005 AAGAGCTGCTGAGCCCTAGCCCAAGAAAACAAAGCTTTCATCTGAGTGTGATGTGCT 1064
DB 298 GInuGInuIleAAsnAlaIleAAsnLySProPheLySMetIleProGlyValIleAAsPVal 317
QY 1065 TTCTTTCTTAATGAGCTCTGATCTATGCTTGTGCAAAAGATTTAAAGCAATTCCTCC 1124
DB 318 PheLeuProAAsnArgHnSProGInuGInuLeuAAsnIleAAsPProAAsPProAAsP 337
QY 1125 ATCATGAGAGTCAATTAACAGAGTGTGCTCTGCTGCT-----ATGAAG 1172
DB 338 ILeValIleGlyAlaAAsnAAsnGInuPheGlyYrIleProIleProLySValMetAArgIleYr 357
QY 1173 GAGGCTCTGAGATCTCAAGTGGCTCAACAAGTCCCTTCCATCTGATACAAAG 1232

```

```

DB 358 AspThrGInuLySglYAlaMetAAsPArgGlyAlaSerGInuAlaAlaLeuGInuLyMetLeuThr 377
QY 1233 ATCTTCACATATCCCGCTCAGATATTGACCTTGTGCTTAATGATTAATCTCCATCAACAAG 1292
DB 378 LeuLeuMetLeuProProThrPheGlyAAsPLeuAAsGInuGlyIleGlyYrAAsPn 397
QY 1293 CACTCCCTCAGTGAATCCGAGACAGAGTCTTCTGAGACTGCTTGGAGATGTGTTCTTGTG 1352
DB 398 GlyAAsPProGInuThrLeuGInuAlaGInPheGInuGInuMetAlaAAsPserMetPheVal 417
QY 1353 GTCCCTGACATGATACAGCTCGATATCAAGAGATGTGTGTCAGACCTGTCTACTTCTAT 1412
DB 418 ILeProAlaLeuGInuAlaHnIlePhe---GInCySerAAsGAlaProValYrPheYr 436
QY 1413 GAGTTTGGGACCCGGGCTCAGTGTGCTTGAACACAGAACCCGCTTTGTCAAGCCGAC 1472
DB 437 GInuPheGInuHnSglInuProSerTyrLeuLySAsnIleArgProProHnIleMetLySAlaAAsP 456
QY 1473 CAGGCTGATGAAGTCCGCTTGTGTGCTGCTGCTGCTGCTCCGAAAGGCGCATTTGATG 1532
DB 457 HnSglYAAsPArgIleuProPheValPhe---ArgSerPhePheGlyYrAAsnYrIleLyS 475
QY 1533 TTCAAGAGAGCCACGAGAGAGAGAAAGTACTAGAGCCGGAAGATGAATACTGGGCT 1592
DB 476 Phe-----ThrGInuGInuGInuGInuGInuLeuSerArgLySMetClySyrTrpAla 492
QY 1593 ACTTTGTCTCAACCGGGAATCTTAATGAGAGAGACCTGTCTGTGCGCAAGTAAAT 1652
DB 493 AAsnPheAlaArgAAsnGlyAAsnProHnSglYLeuProHnItrProLeuPheAAsP 512
QY 1653 CTGACTAGACAGTACTCTCCAGCTGAGACTTGAACATGAGACTCGAGACAGAGATCAAGAA 1712
DB 513 GInuGInuGInuYrLeuGInuGInuAAsnLeuGInuProAlaValGlyYrArgAlaLeuLySAla 532
QY 1713 CCGGAGGTGATTTTGGACCAACACATCCCC 1745
DB 533 HnArgLeuGInuHnThrItrPlySlyAlaLeuPro 543

```

```

RESULT 4
US-09-949-016-9670
US-Sequence 9670, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9670
LENGTH: 577
TYPE: PRT
ORGANISM: Human
US-09-949-016-9670
Alignment Scores:
Pred. No.: 2.76e-111
Score: 1230.00
Length: 577
Percent Similarity: 63.84%
Matches: 248
Beet Local Similarity: 46.70%
Mismatch: 91
Query Match: 31.77%
Indels: 168
DB: 4 Gaps: 24
US-10-023-515-1 (1-2158) x US-09-949-016-9670 (1-577)

```



```

Oy 186 GGGCTTGTGAGAGGCGACAGAGAAACACAGAGCTGGATGATTACAGGCAAGCA 245
Db 44 GlylnapSerAlaSerProIleArgThrThrIleThrGlyGlnValLeuGlySerIleu 63
Oy 246 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTCTTGTGAAGTCCCTTGTCTGT 305
Db 64 ValHleValIleGlyAlaIleAlaIleGlyValGlnThrPheLeuGlyIleProPheAlaIle 83
Oy 306 CCCCCGTGGATCCCTGTGGATTTACGAACCCGAGCCGTGATCCCTCGGATTAACCTTG 365
Db 84 ProProLeuGlyProLeuAlaPheAlaProProGluProProGluSerIleProGlyVal 103
Oy 366 CGAAGAGCCCTTACCTTAACTTGTGCTCCAGAGC-----TCAGAG 410
Db 104 ArgAlaGlyThrThrIleProAlaMetCysLeuGlnAlaPheLeuThrAlaValGluSerGlu 123
Oy 411 TGGCTGTCTTAGATCAACACATGCTCAAGGTGATTAACCCGAAATTCGGATGTCAAGA 470
Db 124 PheIleu-----SerGlnPheAlaMetThrPheProSerAlaSerMetSerGlu 139
Oy 471 GATGCTCTTACCTGAACATCTATGCGCTGCGCCACCGCATACAGCTCCAGCTCCGC 530
Db 140 AspCysLeuGlyLeuSerIleGlyThrProAlaIleSerHleGlyIleSerAlaLeuPro 159
Oy 531 GCTTGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGGCTGACCTGCATCTTGATGGG 590
Db 160 ValMetValIleProIleHleGlyAlaLeuValPheGlyMetAlaSerLeuGlyArgGly 179
Oy 591 TCCGCGCTGTGCTTATAGAGACCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 650
Db 180 SerMetLeuAlaAlaLeuGlnAlaValValValIleIleGlnIleArgLeuGlyVal 199
Oy 651 TTTGGTTTCTTCAACCACTGGAGTACAGATGCTCCGGGAACTGGGACCTTCAAGACAG 710
Db 200 LeuGlyPhePheSerThrGlyAspIleHleAlaThrGlyAsnIleProGlyIleLeuAla 219
Oy 711 GTGGCTGTGTCTGTCTGTGCTCAGAGAAACATCGAGTCTTCCGGTGGAGACCCGAGCT 770
Db 220 ValAlaAlaLeuAlaGlyIleValGlnIleAlaIleAlaIleAlaIleAlaIleAlaIle 239
Oy 771 GTGACCATCTTTGGAGAGTCCGCGAGACCATAGTGTCTTGTGTGTGTGTGTGTGTGT 830
Db 240 ValThrIlePheGlyGluSerIleGlyIleThrSerValSerSerLeuValValSerPro 259
Oy 831 ATGCGCAAGAGCTTATTCACAAAGCCATCATGAGAGTGGGCTGACATCCCTTAC 890
Db 260 IleSerGlnGlyLeuPheHleGlyAlaIleMetGluSerGlyValAlaLeuLeuProGly 279
Oy 891 CTGAGAGCCCATGATTATGAGAGAGTGAAGCTG-----CAGGTGTTGCAATTTTC 944
Db 280 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaIleAlaIle 295
Oy 945 TGTGTGTAACAATGCTCAGACTCTGAGAGCCCTGAGTGGCTGAGAGCAAAACCTTCC 1004
Db 296 SerAlaCysAlaProGlnValAlaPheSerIleValAlaLeuValGlyCysLeuAlaGlyIle 315
Oy 1005 AAGAGCTGTGAGACCTTCAAGCAAAACAAAGCTTTTCACTGAGTGGTGTGAGTGT 1064
Db 316 GluGluIleLeuAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 335
Oy 1065 TTTCTTTCTTAATGAGCTCTAGATCTATTTGTCAAAAAGCTTTAAAGCATTTCTTCC 1124
Db 336 PheIleuProAlaGlnAlaPheGlnIleLeuLeuAlaSerAlaAspPheGlnProValProSer 355
Oy 1125 ATCATCGAGTCAATAACCAAGAGTGTGCTTCTGCTGCTC-----ATGAAG 1172
Db 356 IleValIleGlyAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 375
Oy 1173 GAGGCTCTGAGATCTTCAAGTGTGCTCAACAAAGCTTCTGCTTCACTGTATGACAAAC 1232
Db 376 AspThrGlnIleGlyIleMetAlaAspArgGluAlaSerGlnAlaIleAlaLeuGlnIleMetLeuThr 395

```

```

Oy 1233 ATCTGCAATCCCGCTCAGATATTGCACTTGTGGCTTAATGAATCTTCATGACAAG 1292
Db 396 LeuIleuMetLeuProProIleThrPheGlyAspLeuAlaGluGluIleGlyIleAlaPheAla 415
Oy 1293 CACTCCCTGATGAAATCCGAGACAGCTTCTTGTGAGCTTGTGAGATGTGTTTGTG 1352
Db 416 GlyAspProGlnThrLeuGlnAlaIleAlaIlePheGlnGluIleMetAlaAspSerMetPheVal 435
Oy 1353 GTCCCTGATGATCAACAGCTGATATCAACAGAGATGGCGGTGAGCTGTCTTACTTCTAT 1412
Db 436 IleProAlaLeuGlnValAlaIleAlaIlePhe---GlnCysSerAlaAlaProValIleThrPheIle 454
Oy 1413 GAGTTTCCGACACCGGCTCAGTGTCTTGAAGAACAAGACCGGCTTTGTCTCAAGCCGAC 1472
Db 455 GluPheGlnHleGlnProSerIlePheLeuYasnIleArgProProHleMetCysAlaAsp 474
Oy 1473 CACGCTGATGAATCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1532
Db 475 HleGlyAspGluLeuProPheValPhe---ArgSerPhePheGlyIleYasnIleIleYs 493
Oy 1533 TTGCAAGAGCCACGAGAGAGAGATTAATCTAGACCCGGAAGATGATGAATATCTGGCT 1592
Db 494 Phe-----ThrGluGluGluIleGluIleuSerAlaGlyIleMetCysIleIleProAla 510
Oy 1593 ACCTTTGTGCAACCGGAAATCTTAATGGAACGACTGTCTGTGTGCGCACTTATAT 1652
Db 511 AsnPheAlaAlaGlnAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 530
Oy 1653 CTGACTGAGACGATCTTCAGCTGATTTGAACATGAGCCTTGGACAGAGACTCAAGAA 1712
Db 531 GlnGluGlnIleGlyLeuGlnIleuAlaIleuAlaIleuAlaIleuAlaIleuAlaIleu 550
Oy 1713 CCGCGGTGATTTTGTGACACGACCATCCCC 1745
Db 551 HleArgLeuGlnPheIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 561

RESULT 5
US-09-595-682B-21
; Sequence 21, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danke, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT FILING DATE: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-595-682B-21

Alignment Scores:
Pred. No.: 1,676-103 Length: 565
Score: 1150.50 Matches: 252
Percent Similarity: 56.97% Conservative: 75
Best Local Similarity: 43.90% Mismatches: 180
Query Match: 29.71% Indels: 67
DB: 4 Gaps: 13

US-10-023-515-1 (1-2158) x US-09-595-682B-21 (1-565)
Oy 129 TGGTCTTTTCTGATTTCTCAGCCCTGTGGAGACAGACAGTGGGAAAACTGGG 168
Db 129 TGGTCTTTTCTGATTTCTCAGCCCTGTGGAGACAGACAGTGGGAAAACTGGG 168

```

```

Db      2 TrpLeuCyAlaIleuAlaIleuAlaSerIleuAlaIaCySThrAlaTrpGlyHis----- 19
QY      189 CCTTCGCTGGAAGGGCCAGAGAAACACGAGCTGGAGATTCAAGGCAAGCAATC 248
Db      20 ProSerAlaPro--ProValAlaPheThrValHisGlyLeValIleuGlyLysPheVal 38
QY      249 ACTGTGCTGGGAAGCCCTGTGCTGGAAGCGTGTCCCTGAGAGTCCCTTGTGCTGCC 308
Db      39 SerLeuGlyGlyPheAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 58
QY      309 CGGCTGGAGTCCCTGCGATTACGAACCCGAGCTGCATGCGCTGGAGATAACTGGCA 368
Db      59 ProLeuGlySerLeuAlaGlyPheAlaProGlnProAlaGlnSerTrpSerHisValLys 78
QY      369 GAAGCACCCTCCATATTTGTGGCTCCAGAACTCAGAGTGGCTGTCTTAATCA 428
Db      79 AenThrTrpSerTrpProPheMetCysSerGlnAspAla-----ValSerGly 94
QY      429 CACATGCTC-----AAGGTGCAATTACCG--AAATTGGAGT 464
Db      95 HisMetLeuSerGlyLeuPheThrAsnAlaGlySerGlnIleProLeuLysPhe----- 112
QY      465 TCAGAAAGCTGCTCTACCTGAAACATCTATGCGCTGCCACCGCCATACAGGCTCCAG 524
Db      113 SerGlnAspCysLeuTrpLeuAlaMetIleTrpAlaAspLeuThrLysAlaGlyArg 132
QY      525 CTCCTCCCTCTGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGAGCTCAGCTCATCTT 584
Db      133 LeuProValMetValTrpIleHisGlyGlyGlyLeuMetValGlyGlyAlaSerThrTrp 152
QY      585 GATGGGCGCCGCTGCGCTATGAGAGAGCTGCTGTGCTGCTCCATGACCGGCTA 644
Db      153 AspGlyLeuAlaLeuSerAlaHisGlnAsnValValValAlaThrIleGlnTrpArgLeu 172
QY      645 GGAATATTGGTTCCTTCAACACATGAGATCAGCATGCGCGGAGCTGGGCTTCAAG 704
Db      173 GlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerArgGlyAsnTrpGlyHisLeu 192
QY      705 GACCAAGTGGCTGCTGTCTGCTGCTGCTCAAGAAACATGAGTCTTCGAGTGGAGCC 764
Db      193 AspGlnValAlaAlaLeuAlaGlyTrpValGlnAspAsnIleAlaAsnPheGlyLysPro 212
QY      765 AGCTGTGACCATCTTGTGGCGAGTCCGCGGAGCCATAGGTTTCTATGTTTACTG 824
Db      213 GlySerValThrIlePheGlyLeuSerAlaGlyGlyGlnSerValSerIleLeuLeuLeu 232
QY      825 TCTCCATGCGCAAGGCTTATTCACAAAGCATATGAGAGTGGGGTGGCCATCATC 884
Db      233 SerProLeuThrLysAsnLeuPheHisArgAlaIleSerGlnSerGlyValAlaLeuLeu 252
QY      885 CCTTACCTGAGAGCCCATGATTATGAGAAAGTGAAGAC-----CTCAGAGTGGTTGCA 938
Db      253 SerSerLeu-----PheArgLysAsnThrLysSerLeuAlaGlnIleAla 268
QY      939 CATTTCTGTGTACATAGCTGCTGAGACTGTGAGCGCTGAGGTGAGTGGAGCAAAA 998
Db      269 IleGlnAlaGlyCysLysThrThrSerAlaValMetValHisCysLeuAlaGlnLys 288
QY      999 CCTTCAGAGAGCTGCTG-----ACCTCAGCCAGAAA----- 1031
Db      289 ThrGlnGlnGlnLeuMetGlnValThrLeuLysMetLysPheMetAlaLeuAspLeuVal 308
QY      1032 -----ACAAAGTCTTTCACCTGAGGTGTTGATGGTCTTCTTCTT 1073
Db      309 GlyAspProLysGlnAsnThrAlaPheLeuThrThrValIleAspGlyValLeuLeuPro 328
QY      1074 AATGAGCTGATGATCTTGTCTCAGAAAGCATTTAAAGCAATCTCTTCATCATCGGA 1133
Db      329 LysAlaProAlaGlnIleLeuAlaGlnLysLysTrpAsnMetLeuProLysMetValGly 348
QY      1134 GTCAATAACACAGAGTGTGCTGCTGCTGCTGCTATG----- 1169
Db      349 IleAsnGlnGlnGlnPheGlyTrpIleIleProMetGlnMetLeuGlyTrpProLeuSer 368

```

```

QY      1170 -----AAGAGGCTCTGAGATCTTCAGTGGCTCCAAAGTCCCTT 1211
Db      369 GlnGlyLysLeuAspGlnLysThrAlaThrGlnLeuLeuTrpLysSer----- 384
QY      1212 GCCCTCATCTGTATCAAAACATCCGACATCCCGCTCAGATATTTCACCTGTGGCT 1271
Db      385 -----TrpProIleValAsnValSerLysGlnLeuThrProValAlaThr 399
QY      1272 AATGAATATCTTCATGACAGACATCCCTGACCTGAATCCGAGACAGTCTTGGACTTG 1331
Db      400 GlnLysTrpLeuGlyGlyThrAspAspProValLysLysLysAspLeuPheLeuAspMet 419
QY      1332 CTGGAGATGTTGCTTGTGGTCCCTGACATGATCAGCTCCGATATCAGAGATGCT 1391
Db      420 LeuAlaAspLeuPhePheGlyValProSerValAsnValAlaAspHisHisArgAspAla 439
QY      1392 GGTGCACTGTCTACTTCTATGAGTTTGGCAGCCGCTCAGTGTCTTGAAGACAGAG 1451
Db      440 GlyAlaProThrTrpMetTrpGlyTrpArgTrpArgProSerPheSerAspMetArg 459
QY      1452 CCGGCTTTTTCAAAGCCGACAGCTGATGAAGTCCGCTTGTGTGGTGGCTTC 1511
Db      460 ProLysThrValIleGlyAspHisGlyAspGlnIlePheSerValLeuGlyAlaProPhe 479
QY      1512 CTGAAGGGGAGCATTTGTTATGTTTGAAGAGCCAGAGAGAGAGAAATTACTGAGCCGG 1571
Db      480 LeuLys-----GlnGlyAlaThrGlnGlnGlnIleLysLeuSerLys 493
QY      1572 AAGATGATGAATATCTGGGCTACTTGTGCTGCAACCGGAATCTTAATGGAACGACTG 1631
Db      494 MetValMetCysTrpTrpAlaAsnPheAlaArgAspGlyAsnProAsnGlyGlnGlyLeu 513
QY      1632 TCTGTGGCGCAGTTTAACTGACTGAGCAGTACCTCAGTACCTGAGTGAACATGAGAC 1691
Db      514 ProGlnTrpProAlaTrpAspTrpLysGlnGlyTrpLeuGlnIleGlyAlaThrTrpGln 533
QY      1692 CTGCAAGAGACTCAAAAGAACCGCGGTGGATTTTGGACC 1733
Db      534 AlaAlaGlnLysLeuLysAspLysGlnValAlaAlaPheTrpThr 547

RESULT 6
US-09-264-737-2
/ Sequence 2, Application US/09264737A
/ Patent No. 6107549
/ GENERAL INFORMATION:
/ APPLICANT: Feng, Paul C.C.
/ APPLICANT: Ruff, Thomas G.
/ TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
/ FILE REFERENCE: 38-21(10551) R1E3 Pyridine Tolerance
/ CURRENT APPLICATION NUMBER: US/09/264,737A
/ EARLIER APPLICATION NUMBER: 60/077,377
/ EARLIER FILING DATE: 1998-03-10
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: Patencin Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 566
/ TYPE: prt
/ ORGANISM: Rabbit
US-09-264-737-2

Alignment Scores:
Pred. No.: 2.5e-102
Score: 1138.50 Length: 566
Percent Similarity: 56.62% Matches: 250
Best Local Similarity: 43.55% Conservative: 75
Query Match: 29.40% Mismatches: 176
DB: 3 Gaps: 73
Indels: 73
Gaps: 14

US-10-023-515-1 (1-2158) x US-09-264-737-2 (1-566)

```


Query Match: 29.20% Indels: 67
 DB: 4 Gaps: 13
 US-10-023-515-1 (1-2158) x US-09-595-682B-26 (1-543)

QY 129 TGGTCCTTTTCCTGATTCTCCAGCCCTGTTGGGACACAGACAGTGGGAAAAACTGGG 188
 DB 2 TTPLeuCyAlaLeuAlaLeuAlaSerLeuAlaAlaCysThrAlaIlePheGlyHis----- 19
 QY 189 CCTTCTGTGAAGGGCCACAGAGGAACCCAGCGCTGGATGATTACAGGGCAAGATC 248
 DB 20 ProSerAlaPro-----ProValAlaAspThrValHisGlyLysValLeuGlyLysPheVal 38
 QY 249 ACTGTCTGGGAAGCCCTGTGCTGTGAACGTGTTCCTGGAGTCCCTTGGCTGCTGCC 308
 DB 39 SerLeuGluGlyPheAlaGlnProValAlaValaPheLeuGlyValProPheAlaLysPro 58
 QY 309 CCGCTGGGATCCCGCATTTACGAACCCGACGCTGCATCGCCCTGGATTAACCTGGCA 368
 DB 59 ProLeuGlySerLeuAlaGlyPheAlaProProGlnProAlaGlnSerTrpSerHisValLys 78
 QY 369 GAAGCCACCTCTCACTTAATTGTGCTCCAGAACCTCAGAGTGGCTCTTAGATCAA 428
 DB 79 AsnThrTrpSerTrpProPromeCysSerGlnAspAla-----ValSerGly 94
 QY 429 CACATGTCT-----AAGGTGCATTACCCG---AAATTCCGAGTG 464
 DB 95 HisMetLeuSerGluLeuPheThrAsnAlaGlyGlnAlaSerLeuLysPhe----- 112
 QY 465 TCAGAAAGACTGCTTACCTGAACATCTATGCGCTCCACCCGATACAGGCTTCAAG 524
 DB 113 SerGlnAspCysLeuTrpLeuAsnIleTrpThrProAlaAspLeuThrLysArgLys 132
 QY 525 CTCCCGCTCTGGTGTGTTCCTCCAGAGAGTGCCTTCAAGACTGGCTCAGCTCCATCTT 584
 DB 133 LeuProValMetValTrpIleHisGlyGlyLeuMetValGlyLysAspThrTrp 152
 QY 585 GATGGTCCGCGCTGCTGCTATGAGACGCTGCTGTGTGTGTGCTCCATGACCGGCTA 644
 DB 153 AspGlyLeuAlaLeuSerAlaHisGlnAsnValValValThrIleGlnLysArgLeu 172
 QY 645 GGAATATTTGTTCTTACACATGAGGATCAGCATGCTCCGGGAACTGGGCTTCAAG 704
 DB 173 GlyIleTrpGlyPhePheSerThrGlyAspGluHisSerArgGlyAsnTrpGlyHisLeu 192
 QY 705 GACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
 DB 193 AspGlnValAlaAlaLeuArgTrpValGlnAspAsnIleAlaAsnPheGlyGlyAspPro 212
 QY 765 AGCTGTGTGACCATCTTTGGCGAGTCCGCGGAGGCATTAAGTGTTCATGCTTAATCTG 824
 DB 213 GlySerValThrIlePheGlyGlnSerAlaGlyGlnSerValSerIleLeuLeuLeu 232
 QY 825 TCTCCCATGGCCAAAGGCTTATTCACAAAGCCATGAGAGTGGGCTGGCCATCATC 884
 DB 233 SerProLeuThrLysAsnLeuPheHisArgAlaIleSerGlnSerGlyValAlaLeuLeu 252
 QY 885 CTTTACTGAGAGGCCCATGATTATGAGAAAGATGAGC-----CTGCAGGTGTTGCA 938
 DB 253 SerSerLeu-----PheArgLysAsnThrLysSerLeuAlaGlyLysIleAla 268
 QY 939 CATTTCTGTGTAAACATGCTCAGACTCTGAGGCCCTGCTGAGGTCCCTGAGGACAA 998
 DB 269 IleGlnAlaGlyCysLysThrThrSerAlaValMetValHisCysLeuAlaGlnLys 288
 QY 999 CCTTCAGAGAGCTGCTG-----ACCTCAGCCAGAA----- 1031
 DB 289 ThrGluGluGluLeuMetGluValThrLeuLysMetLysPheMetAlaLeuAspLeuVal 308
 QY 1032 -----ACAAAGCTTTTCACTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1073
 DB 309 GlyAspProLysGluAsnThrAlaPheLeuThrThrValIleAspGlyValLeuLeuPro 328

QY 1074 AATGACCTCTAGATCTATGTCTCGAAAAGCATTTAAAGCAATTCTTCATCATCGGA 1133
 DB 329 LysAlaProAlaGlnIleLeuAlaGluLysLysTrpAsnMetLeuProTrpMetValGly 348
 QY 1134 GTCAATTAACCAAGAGTGGCTTCTGCTGCTCATG----- 1169
 DB 349 IleAsnGlnGlnLysPheGlyTrpIleIleProMetGlnMetLeuGlyTrpProLeuSer 368
 QY 1170 -----AAGAGGCTCTGAGATCCCATGCTGCTCACAACATCCCTT 1211
 DB 369 GlnGlyLysLeuAspGlnLysThrAlaThrGlnLeuLeuTrpLysSer----- 384
 QY 1212 GCCCTCCATCTGATTAACAAACATCTCGACATCCCGCTCAGTATTTGCACTTGTGGCT 1271
 DB 385 -----TrpProIleValAlaAsnValSerLysGlnLeuThrProValAlaThr 399
 QY 1272 AATGAATCTTCCATGACAAACACATCCCTGACTGTAATAATCCGACAGCTTCTGGACTTG 1331
 DB 400 GlnLysTrpLeuGlyGlyThrAspAspProValLysLysLysAspLeuPheLeuAspMet 419
 QY 1332 CTGGAAGATGTTCTTGTGTGCTCCGACATGATCAGACTGATGATGATGATGATGATGATGAT 1391
 DB 420 LeuAlaAspLeuLeuPheGlyValProSerValAsnValAlaArgHisIleAspAla 439
 QY 1392 GATGACCTGTCTACTTCTTATGAGTTTCCGACCGGCTTCAAGCTTTTGAAGACACGAAG 1451
 DB 440 GlyAlaProThrTrpMetTrpGlyTrpArgTrpArgProSerPheSerSerAspMetArg 459
 QY 1452 CCGGCTTTTGTCAAAAGCCGACACGCTGATGAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCT 1511
 DB 460 ProLysThrValIleGlyAspHisGlyAspGluIlePheSerValLeuGlyAlaProPhe 479
 QY 1512 CTGAAGGGGACATTTGATTGTTTCAAGAGACCAAGAGAGAGAGAACTTACAGCCGG 1571
 DB 480 LeuLys-----GlnGlyAlaThrGlnGlnGlnIleLysLeuSerLys 493
 QY 1572 AAGATGATGAATATCTGGGCTACTTGTGCTCGAACCGGGAATCTTAATGGAAACGACTG 1631
 DB 494 MetValMetLysTrpTrpAlaAsnPheAlaArgAsnGlyAspProAsnIleGlnLysLeu 513
 QY 1632 TCTCTGTGGCCAGCTTATTAATCTGATGAGCAGTACTTCACGCTGACTTGAATGAGC 1691
 DB 514 ProGlnTrpProAlaTrpAspTrpLysGlnGlyTrpLeuGlnIleGlyAlaThrThrGln 533
 QY 1692 CTCGACAGAGACTCAAAAGAACCGCGGCTG 1721
 DB 534 AlaAlaGlnLysLeuLysAspLysGlnVal 543

RESULT 8
 US-08-845-295A-2
 / Sequence 2, Application US/0845295A
 / Patent No. 5817490
 / GENERAL INFORMATION:
 / APPLICANT: Hubbe, John C.
 / TITLE OF INVENTION: Enzymatic Process for the Manufacture of
 / TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
 / NUMBER OF INVENTIONS: 2-Keto-L-Gulonic Acid
 / NUMBER OF SEQUENCES: 3
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Eastman Chemical Company
 / STREET: P.O. Box 511
 / CITY: Kingsport
 / STATE: Tennessee
 / COUNTRY: USA
 / ZIP: 37662-5075
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 3.5 inch disk
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM: Windows 95
 / SOFTWARE: Microsoft Word
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/845,295A
 / FILING DATE: 25-April-97

CLASSIFICATION: 435
 PRIOR APPLICATION DATA: US 60/017,879
 APPLICATION NUMBER: 17-MAY-1996
 FILING DATE: 17-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Cheryl J. Tubach
 REGISTRATION NUMBER: 38,346
 REFERENCE/DOCKET NUMBER: 70432
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 423-229-6189
 TELEFAX: 423-229-1239
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 US-08-845-295A-2

Alignment Scores:

Pred. No.:	1,02e-97	Length:	584
Score:	1091.50	Matches:	259
Percent Similarity:	54.39%	Conservative:	63
Best Local Similarity:	43.75%	Mismatches:	191
Query Match:	28.19%	Indels:	79
DB:	2	Gaps:	17

US-10-023-515-1 (1-2158) x US-08-845-295A-2 (1-584)

```

QY 129 TGGTCTTTTCTGATATCTCAGCCCTGTGGGACACAGATGGGAAAACTGGG 188
DB 2 TrrleuLeuProleuValLeuThrSerleuAlaSerSerAlaThrTrp-----AlaGly 19
QY 189 CTTTCTCTGAAGGGCCACAGAGAAACACAGGCTGGATTCAGGCGAAGCAAGTC 248
DB 20 GlnProAlaSerProProValValAspThrAlaGlnGlyArgValLeuGlyIyStrVal 39
QY 249 ACTGTG-----CTGGGAAGCCCTGTGCTGGGAACGTTCTGGAGTCCCTTTGCT 302
DB 40 SerLeuGlnGlyLeuAlaPheThrGlnProValAlaValPheLeuGlyAlaProPheAla 59
QY 303 GCTCCCGCGCTGGATCCCTGGATTTACGAACCCGCGACCTGCATCGCCTGGGATTAAC 362
DB 60 LysProProLeuGlnGlySerleuArgPheAlaProProGlnProAlaGlnIuproTrpSerPhe 79
QY 363 TTGCGAAGCAACCTCTTACCTTAATTGTGCTTCAAGACTCA-----GAGTGGCTG 416
DB 80 ValLysAsnThrThrSerTyProProMetCysGlyGlnAspProValValGlnGlnMet 99
QY 417 CTCTTAGAT-----CAACACATGCTCAAGGTGCATTACCCGGA 455
DB 100 ThrSerAspLeuPheThrAsnPheThrGlyLysGlnArgLeuThrLeuGlnPhe----- 117
QY 456 TTCGAGTGTCAAGAGACTGCTCTTACCTGAACATCTATGCGCTGCCACCGCATACA 515
DB 118 -----SerGlnAspCysLeuTyTrleuAsnIleTyThrProAlaAspLeuThrLys 134
QY 516 GCGTCCAAAGCTCCCGCTTGTGTGGTTCCAGAGAGTGCCTTCAAGACTGAGTCAAGC 575
DB 135 ArgGlyArgLeuProValMetValTrpIleAsGlyGlyGlyLeuValLeuGlyGlyAla 154
QY 576 TGCATCTTTGATGGGCGCGCTGGCTGCTATGAGAAC-----GTGCTGTTGTGGTC 629
DB 155 ProMetTyArgPglValValValLeuAlaValAsnGlnAsnPheThrValValValAla 174
QY 630 GTCCAGTACCGGCTGAGATATTTGGTTCTTCAACACATGGAGATCAGATCTCCGGG 689
DB 175 IleglnTyArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGlnIleSerArgGly 194
QY 690 AACTGGGCTTCAAGACCAAGGCTGCTGTCTGTCTGGGCTCAAGAAACATCGAGTTC 749
DB 195 AsnTrpGlnIleLeuAspGlnValAlaAlaLeuIleTrpValGlnGlnIleAlaAsn 214

```

```

QY 750 TTCGTGGGAGCCCGGAGCTCTGTGACCATCTTTGGCGAGTCC-----GCGGAGCCATA 803
DB 215 PheGlyGlyAspProGlySerValThrIlePheGlyGlySerPheThrAlaGlyGlyIu 234
QY 804 AGTGTTCATGCTTAACTGTCTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATG 863
DB 235 SerValSerValLeuValLeuSerProLeuAlaLysAsnLeuPheIleArgAlaIleSer 254
QY 864 GAGAGTGGGGTGGCCATCCCTTACCTGGAGGGCCCATGATTAATGAGAAAGTGAGAC 923
DB 255 GluSerGlyValAlaLeuThrValAlaLeuValArgLysAspMetLysAlaAlaLys 274
QY 924 CTCAGTGTGTGCACATTTCTGTGTGTACAAATGGCTCAGACTGAGGCGCTCGAGG 983
DB 275 GlnIleAlaValLeuAlaGlyCysLysThrThrThrSerAlaValAlaPheThrPheValHis 294
QY 984 TGCCTGAGGACAAACCTCCCAAGAGCTGTGACCTCAGCCAGAAACAAAGCTTTTC 1043
DB 295 CysLeuArgGlnLysSerGlyAspGlyLeuLeuAspLeuThrLeuLysMetLysPheLeu 314
QY 1044 ACT-----CGAGTGGCTGAT 1058
DB 315 ThrLeuAspPheIleGlyAspGlnArgGlyLysSerHisProPheLeuProThrValValAsp 334
QY 1059 GGTGCTTTCTTCTTAATGAGCTCTAGATCTATTGTCTCAGAAA-----GCATTTAA 1112
DB 335 GlyValLeuLeuProLysMetProGlnGlnIleLeuAlaGlnLysAspPheThrPheAsn 354
QY 1113 GCAATTCCTTCCATCATCGGAGTCAATTAACAGAGTGTGCTTCTGTGCTCT----- 1166
DB 355 ThrValProTyIleValGlyIleAsnLysGlnGlnPheGlyTyPheLeuProThrMet 374
QY 1167 -----ATGAGGAGGCTCTCAGATCCCTCAT 1193
DB 375 MetGlyPheProLeuSerGlnGlyLysLeuAspGlnLysThrAlaThrSerLeuLeuTrp 394
QY 1194 GCGTCCAAAGTCCCTTCCCTCATCTGATCAAAACATCTGCACATCCCGCTCAG 1253
DB 395 LysSer-----TyrProIleAlaAsnIle-----ProGlu 404
QY 1254 TATTGACCTTGTGGCTAATGATTAATCTTCATGACAGAC-----TCC 1298
DB 405 GluLeuThrProValAlaThr-----PheThrAspLysTyLeuGlyGlyThrAspAsp 422
QY 1299 CTGACTGAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTCTTTGTGCTCCT 1358
DB 423 ProValLysLysAspLeuPheLeuAspLeuMetGlyAspValValPheGlyValPro 442
QY 1359 GCAGTGAATCAGCTCGATATGACAGAGATGCTGTGACCTGTCTACTTATGAGTT 1418
DB 443 SerValThrValAlaArgGlnHisArgAspAlaGlyAlaProThrTyMetTyGlnPhe 462
QY 1419 CGGACCGGCTCAAGTGTCTTGAGAC-----ACGAACCGGCTTTGTCAAGCCGAC 1472
DB 463 GlnTyArgProSerPheSerSerAspLysPheThrLysProLysThrValIleGlyAsp 482
QY 1473 CACGCTGATGAAGTCCGCTTGTGTGCTGCTGCTCTTCAAGGGGACATTTGATG 1532
DB 483 HisGlyAspGlnIlePheSerValPheGlyPheProLeuLeuLysGlyAsp----- 499
QY 1533 TTCGAAGGACCAAGGAGGAGAAAGTAACTTACGCGGAGATGATGAATAGTGGGT 1592
DB 500 -----AlaProGlnGlnGlnValSerLeuSerLysThrValMetLysPheThrAla 516
QY 1593 ACCTTGTCTGAAACCGGGAATCTTAATGGAACGACTGTCTGTGGCA-----GCT 1646
DB 517 AsnPheAlaArgSerGlyAsnProAsnGlyGlyLeuProHisIleTrpProPheThrMet 536
QY 1647 TATAATCTGACTGACAGTACTCAGCTGCACTTGAACATGAGCTCGAGACAGACTTC 1706
DB 537 TyAspArgGlnGlnGlyTyIleGlnIleIleGlyValAsnThrGlnAlaAlaLysArgLeu 556
QY 1707 AAAGAACCGGCGGTGATTTTTCAGCAGCAGCAGC 1742

```

```

Db      557  LysGlyGluGluValAlaIleThrPheAspLeu 568
RESULT 9
/ Sequence 2, Application US/09140933
/ Patent No. 6022719
/ GENERAL INFORMATION:
/ APPLICANT: Hubbs, John C.
/ TITLE OF INVENTION: Enzymatic Process for the Manufacture of
/ TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
/ TITLE OF INVENTION: 2-Keto-L-Gulonic Acid
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eastman Chemical Company
/ STREET: P. O. Box 511
/ CITY: Kingsport
/ STATE: Tennessee
/ COUNTRY: USA
/ ZIP: 37662-5075
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch disk
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows 95
/ SOFTWARE: Microsoft Word
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/140,933
/ FILING DATE: 27-August-98
/ CLASSIFICATION:
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 60/017,879; 08/845,295
/ FILING DATE: 17-May-96; 25-April-97
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Cheryl J. Tudach
/ REGISTRATION NUMBER: 38,346
/ REFERENCE/DOCKET NUMBER: 70432
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 423-229-6189
/ TELEFAX: 423-229-1239
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 584 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ MOLECULE TYPE: protein
/ US-09-140-933-2

Alignment Scores:
Pred. No.: 1.02e-97 Length: 584
Score: 1091.50 Matches: 259
Percent Similarity: 54.39% Conservative: 63
Best Local Similarity: 43.75% Mismatches: 191
Query Match: 28.19% Indels: 79
Gaps: 17

US-10-023-515-1 (1-2158) x US-09-140-933-2 (1-584)
QY      129  TGGTCTTTTCTGATTCCTCAGCCCTCTGGACACACAGCTGGGAAAACTGGG 188
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2   TTrpleuLeuProIeValIeValIeThSerIeValIeThrPhe-----AlaGly 19
QY      189  CCTTCTGCTGAAGGCCACAGAGAACACAGAGCTGGATGATTACGGGCAAGCAAGTC 248
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      20  GlnProIaIeSerProIeValIeValIeThrAlaGlnGlyArgValIeGlyLysTrVal 39
QY      249  ACTGG-----CTGGGAAGCCCTGTGCTGTGAACGTGTTCTCGAGAGTCCCTTTGCT 302
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      40  SerIeugIuGlyLeuAlaPheThGlnProValAlaValPheIeugIyValIProPheAla 59
QY      303  GCTGCTCCCGCTGGGATCCCTGCGATTACGAACCCGAGCTGCATGCGCCCTGGGANAAC 362
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      60  LysProIeProIeugIySerIeValIePheAlaIeProIeGlnProIaIeGluProIeTrPhe 79

```

```

QY      363  TTGCGAAGAACCACTCTACCTAATTGTCCTCCAGAACTCA-----GAGTGGCTG 416
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      80  ValIyAsnThrThSerIyProIePheCysGlnAspProValIeGluIeGln 99
QY      417  CTCTTGAT-----CAACATGCTCCAGTGCATTAACCCGAAA 455
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      100  ThSerAspLeuPheThrAsnPheThrGlyLysGluArgLeuThrLeuGluPhe----- 117
QY      456  TTGCGAGTGTCAAGAGCTGCTACCTACCTACCTATGCGCCCTCCAGCCGATCA 515
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      118  -----SerIeAspCysLeuIyTrIeAsnIleIyTrIProIaIeAspLeuThrLys 134
QY      516  GGCTCCAGCTCCCGCTGTGCTGTGCTCCAGAGGCTGCTCAAGACTGCTGAGCC 575
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      135  ArgGlyArgLeuProValIeMetValIeTrIleIeGlyGlyLeuValIeGlyAla 154
QY      576  TCCATCTTGATGGGTCCGCTGCTGCTGCTGCTATAGAC-----GTGCTGTTGGTGC 629
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      155  ProMetIyTrAspGlyValIeValIeValIeAlaIeIleGluAsnPheThrValIeValIa 174
QY      630  GTCCAGTACCGGCTAGAAATATTTGTTTCTTCCACATGAGTACGATGCTCCGGGG 689
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      175  IleGlnIyTrArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGluIeSerArgGly 194
QY      690  AACTGGGCTTCAGAGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      195  AsnTrpGlyIleIeAspGlnValIeAlaIeAlaIeAsnIleTrpValIeGlnIeAlaIeAsn 214
QY      750  TTGCGTGGGAGCCCACTGCTGATCCTTTGGGAGCTC-----GGGGAGGCATA 803
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      215  PheGlyIyAspProIeGlySerValIeThrIlePheGlyIeIeSerPheThrAlaGlyGly 234
QY      804  AGTGTTCCTGACTTATTAATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATG 863
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      235  SerValIeValIeValIeValIeSerProIeValIeAlaIeAsnIlePheIleIeSer 254
QY      864  GAGAGTGGGAGTGGCCATCATCCCTTACCTGAGAGCCCATGATATATAGAAAGCTGAGAC 923
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      255  GluSerIyValAlaIeIeThrValAlaIeValAlaIeArgLysAspMetLysAlaIeAlaLys 274
QY      924  CTGCAAGTGTGCAATTTCTGTGTGAACATGCTGACACTGAGCCCTGCTGAGG 983
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      275  GlnIleAlaValIeValIeGlyCysLysThrThrThrThrSerAlaValIePheThrPheValIe 294
QY      984  TGCCTGAGCAAAACCCCTCCAGAGCTGCTGACCTCAGCCGAAAAAAGTCTTTC 1043
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      295  CysIeugIyGlnIySerGluAspGluIeLeuAspLeuThrLeuIyMetLysPheLeu 314
QY      1044  ACT-----CGAGTGGTGAAT 1058
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      315  ThrIeAspPheIleGlyAspGlnArgGluIeSerIleProIePheLeuProIeValIeAsp 334
QY      1059  GGTGCTTCTTCTTCTTAATAGCTCTAGATCTATTGTCTCAGAA-----GCATTTAA 1112
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      335  GlyValIeLeuLeuProIySweCProIeGluIleIeAlaIeGluLysAspPheThrPheAsn 354
QY      1113  GCAATCTCTTCCATGAGTCAATTAACACAGAGTGGGCTGCTGCTGCT----- 1166
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      355  ThrValProIyTrIleValIeGlyIleAsnIyGlnIeIeGlyTrIlePheLeuProIeThrMet 374
QY      1167  -----ATGAAGAGGCTCTGAGATCCTCAGT 1193
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      375  MetGlyPheProIeSerGluIyLysIeAspGlnLysTrIleAlaThrSerIeLeuTrp 394
QY      1194  GGCTCCACAAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      395  LysSer-----TyrProIleAlaIeIle-----ProIeL 404
QY      1254  TATTTCACCTTGTGCTAATGAATATCTTCATGACACAGC-----TCC 1298
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      405  GluIeIeThrProValIeAlaIeThr-----PheThrAspLysTrIleGlyIyThrAspAsp 422
QY      1299  CTGACTGAATCCAGACAGACTCTTCTGACTTGTGAGATGTCTTGTGCTGCTCT 1358

```

Db	423	ProValIysbIysbIysbAspLeuPheLeuaspLeuMetGlyAspValValPheGlyValPro	442
Qy	1359	GCACGTGATCAGAGCTCGATATCACAGAGATCTGTGTGACACCTGTCTACTTATGAGTTT	1418
Db	443	SerValThrValAlaIaGlnIleAspGlyAspIleGlyAlaProThrTyrMetCysIuIuPhe	462
Qy	1419	CGGACCGGGCTCAGTGTCTTGAAGAC-----ACGAAGCGGGCTTTGTTCMAAGCCGAC	1477
Db	463	GlnTyrArgProSerPheSerSerAspIysAspPheThrIlysProLysThrValIleGlyAsp	482
Qy	1473	CAGCGTATGAGTCCGCTTGTGTGTGCGTGGTGCGCTTCCGAAGGGGAGCATTTGTTATG	1532
Db	483	HisGlyAspGlyIuIlePheSerValPheGlyPheProLeuIleuLysGlyAsp-----	499
Qy	1533	TTCGAAGAGCCACGGAGAGAGAAGTTATCTGAGCCGAGAGATGATGAATTAATCTAGGCT	1592
Db	500	-----AlaProGlnGluGluValSerLeuSerLysThrValMetLysPheThrPala	516
Qy	1593	ACCTTGTCTCGAACCGGAAATCTTAATGGAGACACCTGTCTGTGGCCA-----GCT	1648
Db	517	AsnPheAlaIaArgSerGlyAsnProAsnGlyIuGlyLeuProHisIleTrpProPheThrMet	536
Qy	1647	TATATATCTGACTGAGCACTACCTCCAGCTCGACTTGAATCGAAGCCTTGGACAGAGATC	1706
Db	537	TyrAspGlnGluGlnGlyTyrLeuGlnIleGlyValAsnThrGlnAlaAlaLysArgLeu	556
Qy	1707	AAAGAACCGGGCGGTGATTTTGGACCGACGACATCC	1742
Db	557	LysGlyGluGluValAlaPheTrpAsnAspLeuLeu	568

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

US-09-146-661-2					
Alignment Scores:					
Pred.	No.:	1,02e-97	Length:	584	
		1091.50	Matches:	259	
Score:		54.39%	Conservative:	63	
Percent Similarity:		43.75%	Mismatches:	191	
Best Local Similarity:		28.19%	Indels:	79	
Query Match:			Gaps:	17	
DB:		3			
US-10-023-515-1 (1-2158) x US-09-146-661-2 (1-584)					

US-10-023-515-1 (1-2158) x US-09-146-661-2 (1-584)

QY 129 TGGGCTTTTTCCTGATTTCTCCAGCCCTCTGTGGGAGACAGACAGTGGGGAAAACTGGG 188
Db 2 TTPLeuLeuProLeuValLeuThrSerLeuLaserSerIalThrTP-----AlaIy 19
QY 189 CCTTCTCTGAGAGGCCACAGAGAGACACAGCGTGGATGATTCAGGGCAAGCAACTC 248
Db 20 GlnProLaserProProValValAlaPThrIalGlnIyAlaGValLleuGlyLlyTyVal 39
QY 249 ACTGTG-----CTGGAAAGCCCTGTGCTGTGAACGTGTTCCTGGAGTCCCTTGTCT 302
Db 40 SerLeuGlnIyLeuAlaPheThrGlnProValAlaValPheLeuGlyValProPheAla 59
QY 303 GCTCCCCCGGTGGGATCCCTGCAGATTTCAGAACCCGAGCGCTGATGCGCTGGGATAC 362
Db 60 LysProLeuGlnGlySerLeuArgPheAlaProProGlnProAlaGlnProIzSerPhe 79
QY 363 TTGGGAAGAGCCACTCTACCCCTAAATTGGTGCCTCCAGAACTCA-----GAGTGGCTG 416
Db 80 ValIyAsnThrThrSerTyProProMetCysGlnAsnProValValGlnGlnMet 99
QY 417 CTCTTAGAT-----CAACACATGCTCAAGTGCATTAACCCGAAA 455
Db 100 ThrSerAspLeuPheThrAsnPheThrGlyLysGlnArgLeuThrLeuGlnPhe----- 117
QY 456 TTGGGAGTGCAGAAAGCTGCTTACTACTGAACCTATAGCGCTGCCACCCGATACA 515
Db 118 -----SerGlnAspCysLeuTyLeuAsnIleTyThrProAlaAspLeuThrLys 134
QY 516 GGCTCAAGGCTCCCGCTTGTTGGTGTGGTCCACAGAGGTGCTTCAGAGCTGGGCTCAACC 575
Db 135 ArgGlyArgLeuProValMetValTrrIleHisGlyGlyLysLeuValLleuGlyValAla 154
QY 576 TCCATCTTTTATGGGTCGCGCTGGCTGGCTATGAGAC-----GTGCTGGTGTGCTC 629
Db 155 ProMetTyTrAspGlyValValLeuAlaAlaHisGlnAsnPheThrValValValAla 174
QY 630 GTCCAGTACCGGCTAGAGATATTGGTTTCTTCAACACATGGATCAGCATGCTCCGGGG 689
Db 175 IleGlnTyTrArgLeuGlyIleTyGlyPhePheSerThrGlyAspGlnHisSerArgIy 194
QY 690 AACCTGGCCCTTCAAGACACAGTGGCTGTCTGTCTCGGTCCAGAAAGATGAGATTC 749
Db 195 AsnTrpIyHisLeuAspGlnValAlaAlaLeuHisTrpValGlnGlnAsnIleAlaAsn 214
QY 750 TTGGGTGGGAGCCACAGCTCTGTGACCATCTTTGGCAGAGTCC-----GGCGAGCCATA 803
Db 215 PheGlyIyAsnProGlySerValThrIlePheGlyGlnSerPheThrIalGlyGly 234
QY 804 AGTGTATTCTAGCTTTATACTGTCTCCCATGGGCCCAAGAGCTTATTCCAAAGCCATCATG 863
Db 235 SerValSerValLeuValLeuSerProLeuAlaLysAsnLeuPheHisArgAlaIleSer 254
QY 864 GAGAGTGGGGTGGCCATCATCCCTTACTCTGAGGCCCATGATTATGAGAAAGTGAAGAC 923
Db 255 GlnSerIyAlaAlaLeuThrValAlaLeuValArgLysAspMetIyAlaAlaAlaLys 274
QY 924 CTGAGAGTGGTGCACATTTCTGTGGTAAACAATGCTCAGACTGTAGAGCCCTGAGAG 983
Db 275 GlnIleAlaValLeuAlaGlyCysIySerThrThrThrSerIalValAlaPheThrPheValHis 294
QY 984 TGCGTAGAGCAAAACCTTCAAGAGAGCTGTGACCTTGACCGCAAAAACAAGCTTTTC 1043

[illegible]

```

1      STATE: Tennessee
2
3      COUNTRY: USA
4
5      ZIP: 37662-5075
6
7      COMPUTER READABLE FORM:
8
9      MEDIUM TYPE: 3.5 inch disk
10
11     COMPUTER: IBM Compatible
12
13     OPERATING SYSTEM: Windows 95
14
15     SOFTWARE: Microsoft Word
16
17     CURRENT APPLICATION DATA:
18
19     APPLICATION NUMBER: US/09/150,515
20
21     FILING DATE: 09-SEP-1998
22
23     CLASSIFICATION:
24
25     PRIOR APPLICATION DATA:
26
27     APPLICATION NUMBER: US 60/017,879; 08/845,255
28
29     FILING DATE: 17-May-96; 25-April-97
30
31     ATTORNEY/AGENT INFORMATION:
32
33     NAME: Cheryl J. Tubach
34
35     REGISTRATION NUMBER: 38,346
36
37     REFERENCE/DOCKET NUMBER: 70432
38
39     TELECOMMUNICATION INFORMATION:
40
41     TELEPHONE: 423-229-6189
42
43     TELEFAX: 423-229-1239
44
45     INFORMATION FOR SEQ ID NO: 2:
46
47     SEQUENCE CHARACTERISTICS:
48
49     LENGTH: 584 amino acids
50
51     TYPE: Amino Acid
52
53     TOPOLOGY: Linear
54
55     MOLECULE TYPE: protein
56
57     US-09-150-515-2

```

Alignment Scores:

Pred. No.:	1.02e-97	Length:	58
Score:	1091.50	Matches:	25
Percent Similarity:	54.39%	Conservative:	25
Best Local Similarity:	43.75%	Mismatches:	19
Query Match:	28.19%	Indels:	79
DB:	3	Gaps:	17

US-10-023-515-1 (1-2158) x US-09-150-515-2 (1-584)

QY	129	TGGGCTTTTTCGATTTCCGAGCCCTCGTTGGGACACAGACAGTGGGAAAACTGGG	189
Db	2	TrypLeuProLeuValIleuThrSerLeuAlaSerSerAlaThrTP-----AlaGly	19
QY	189	CCTTCCTCTAAGGCCACAGAGAAACACAGCGCTGGAGTGAATTCAAGGCGCACAGTCC	248
Db	20	GlnProAlaSerProProValAlaValAspThrAlaGlnGlyArgValLeuGlyLysTyrVal	39
QY	249	ACTGTG-----CTGGGAAGCCCTGGCGCTGGAGCGAGTTCTCCGGAGTCCCTTTGCT	302
Db	40	SerLeuGlnIuGlyLeuAlaPheThrGlnProValAlaValPheLeuGlyValProPheAla	59
QY	303	GCCTCCCCCGTGGGATCCTCGCATTTACGAACCCGACGCTGCATGCCCTGGAGTAAC	362
Db	60	LysProProLeuGlySerLeuArgPheAlaProProGlnProAlaGlnProTyrSerPhe	79
QY	363	TTGCGAAGAACCACTCTCACTAATTTGGTCCCGACACATCA-----GAGTGGCTG	416
Db	80	ValIysAsnThrThrSerTyrProProMetCysCysGlnAspProValValGlnGlnMet	99
QY	417	CTCTTGAT-----CACACATGTCTCAAGTGCAATCCCGAA	455
Db	100	ThrSerAspLeuPheThrAsnPheThrGlyLysGlnIuArgLeuThrIreuGlnPhe-----	117
QY	456	TTGCGAGTTCAGAGATCGCTCACTAATCACTTAAGCGCTGCCACGCCCATAC	515
Db	118	-----SerGlnAspCysLeuIuTyrLeuAsnIleTyrThrProAlaAspLeuThrIlys	134
QY	516	GGCTCAAGCTCCCGCTTGGTGGTGTCCAGAGAGTGCTTCAAGATCGCTACGCC	575
Db	135	ArgGlyArgLeuProValMetValTyrIleHisGlyGlyGlyLeuValLeuGlyGlyAla	154
QY	576	TCCATCTTTGATGGGTCCGCCCTGGCTGCCTATAGAGAC-----GTGCTGGTTGTGGCT	629


```

Db      155  PrometTyrAspGlyValValLeuAlaIaIaHisGluAsnPheThrValValValAla 174
      630  GTCCAGTACCCGGCTGGAATATTTGGTTTCTCACCAATGGAGATGAGATCTCCGGG 689
      175  IIEGINTYrAArgLeuGlyLettRgLyPhePheSerThrGlyAspGluHiserArgGly 194
      690  AACTGGGCTTCAGACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
      195  AenTrpGlyHisLeuAspGlnValAlaAlaLeuHisIstRValGlnGlnAsnIleAlaAsn 214
      750  TTTCGTTGGGAGACCCAGCTCTGTACCATCTTTGGGAGCTCC-----GGCGAGCCATA 803
      215  PheGlyGlyAspProGlySerValThrIlePheGlyGlySerPheThrAlaGlyGlyGlu 234
      804  AGTGTTCCTAGCTATATCTGCTCCCAAGGCAAGGCTTTTCCAGAAACCATCATG 863
      235  SerValSerValLeuValLeuSerProLeuAlaLysAsnLeuPheHisIstRValIleSer 254
      864  GAGAGTGGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGATTATGAGAGAGTGAAG 923
      255  GlnSerGlyValAlaLeuThrValAlaLeuValArgLysAspMetLysAlaAlaLys 274
      924  CTGCAAGTGTGTCACATTTCTGTGGTAACAATGCGGTCAAGCTGAGGCGCTGCTGAG 983
      275  GlnIleAlaValLeuAlaGlyCysLysThrThrThrSerAlaValAlaPheThrValHis 294
      984  TGCCTGAGAGCAAAACCTCCCAAGAGCTGAGACCTTCAGCAGAAACAAAGCTTTC 1043
      295  CysLeuArgGlnLysSerGluAspGlyLeuLeuAspLeuThrLeuLysMetLysPheLeu 314
      1044  ACT-----CGAGTGGTGTAT 1058
      315  ThrLeuAspPheHisGlyAspGlnArgGluSerHisProPheLeuProThrValAlaAsp 334
      1059  GGTGCTTCTTCTTAATGAGCTCTAGATCTATTGTCTCAGAAA-----GCATTAA 1112
      335  GlyValLeuLeuProLysMetProGluGlnIleLeuAlaGlyLysAspPheThrPheAsn 354
      1113  GCAATTCCTTCATCATGAGAGTCAATAACAGAGTGGCTTCTGCTGCT----- 1166
      355  ThrValProTyrIleValGlyIleAsnLysGlnLysPheGlyTyrPheLeuProThrMet 374
      1167  -----ATGAGAGAGCTCCTGAGATCCCTGACT 1193
      375  MetGlyPheProLeuSerGlnGlyLysLeuAspGlnLysThrAlaThrSerLeuLeuTrp 394
      1194  GAGCTCAACAGTCCCTGCTCCATCTGATACAAACATCTGACATCCCGCTCAG 1253
      395  LysSer-----TyrProIleAlaAsnIle-----ProGlu 404
      1254  TATTTCACCTTGTGGCTAATGATCTTCATGACAGAC-----TCC 1298
      405  GluLeuThrProValAlaIaThr-----PheThrAspLysTyrLysGlyGlyThrAspAsp 422
      1299  CTGATGTAATCCGAGACAGTCTTGGACTGATGATGATGATGATGATGATGATGATGAT 1358
      423  ProValLysLysLysAspLeuPheLeuAspMetGlyAspValAlaPheIleValPro 442
      1359  GCATGATACACAGCTGATATACAGAGATGCTGGTGCACCTGTCTATCTATGAGTTT 1418
      443  SerValThrValAlaArgGlnHisArgAspAlaGlyAlaProThrTyrMetCysGluPhe 462
      1419  CGGACCGGCTCAGTGTCTTGAAGAC-----ACGAGCGGCTTTGTCAAAGCCAG 1472
      463  GlnTyrArgProSerPheSerSerAspLysPheThrLysProLysThrValIleGlyAsp 482
      1473  CACGCTGATGAGTCCGCTTGTGTGCTGGTGGCTTCGTAAGGGAGCATTTGTATG 1532
      483  HisGlyAspGlnIlePheSerValPheGlyPheProLeuLeuLysGlyAsp----- 499
      1533  TTCGAGAGCCACGAGAGAGAGATTACTGAGCCGAGAGATGATGATAATATCTGAGCT 1592

```

```

Db      500  -----AlaProGlnGlnGluValSerLeuSerLysThrValMetLysPheThrAla 516
      1593  ACCTTGTCTGCAACCGGGAATCTTAATGGAAACGACTGTCTCTGGCCA-----GCT 1646
      517  AsnPheAlaArgSerGlyAsnProAsnGlnGlyLysLeuProHisIstRProPheThrMet 536
      1647  TATAATCTGACTGACAGTACTCCAGTCTGAGCTTGAACATGAGCCTCGACAGAGACTC 1706
      537  TyrAspGlnGlnGlnGlyTyrLeuGlnIleGlyValAsnThrGlnAlaAlaLysArgLeu 556
      1707  AAGAACCGCGGTGATTTTGGACAGACCATC 1742
      557  LysGlyGlnGluValAlaPheThrAsnAspLeuLeu 568

RESULT 12
US-09-264-737-1
/ Sequence 1, Application US/09264737A
/ Patent No. 6107549
/ GENERAL INFORMATION:
/ APPLICANT: Feng, Paul C.C.
/ APPLICANT: Ruff, Thomas G.
/ TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
/ FILE REFERENCE: 38-21(1051) RLS3 Pyridine Tolerance
/ CURRENT FILING DATE: 1999-03-09
/ EARLIER APPLICATION NUMBER: US/09/264,737A
/ EARLIER FILING DATE: 1998-03-10
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 539
/ TYPE: PRT
/ ORGANISM: Rabbit
US-09-264-737-1

Alignment Scores:
Pred. No.: 3,73e-92 Length: 539
Score: 1034.50 Matches: 234
Percent Similarity: 55.05% Conservative: 71
Best Local Similarity: 42.24% Mismatches: 176
Query Match: 26,72% Indels: 73
DB: 3 Gaps: 13

US-10-023-515-1 (1-2158) x US-09-264-737-1 (1-539)
      189  CCTTCTGTGAAGGCGCAAGAGAAACCAAGCTGGAGTGGATTCAAGGCAAGCAAGTCT 248
      2  ProSerAlaPro---ProValValAspThrValLysGlyLysValLeuGlyLysPheVal 20
      249  ACTGTGCTGGAGAGCCCTGCTGCTGTAACGTTTCCCTGGAGTCCCTTGTGCTGCTCC 308
      21  SerLeuGlnGlyPheAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 40
      309  CGGCTGGAGATCCCTGCGATTACGAAACCGGAGCTGATGAGCTGCGCTGGATTAATCGCA 368
      41  ProLeuGlnSerLeuArgPheAlaProProGlnProAlaGlnSerTyrPheHisValLys 60
      369  GAAAGCACTTCTTAACCTTAATTTGTGCTCGAAGATCTGAGATGAGTCTCTTATGATCAA 428
      61  AsnThrThrSerTyrProPrometCysSerSerAla-----ValSerGly 76
      429  CACATGCTC-----AAGGTCATTACCG--AAATTGCGAGTG 464
      77  HisMetLeuSerGlnLeuPheThrAsnArgLysGlnAsnIleProLeuLysPhe----- 94
      465  TCAGAAAGCTGCTCTTACTGTAACATCTATGCGCTGCGCAAGCGGATACAGAGCTCAAG 524
      95  SerGlnAspCysLeuTyrLeuAsnIleTyrThrProAlaAspLeuThrLysArgGlyArg 114
      525  CTCCCGGCTTGTGTGTGCTTCCAGAGAGTGCCTTCAAGACTGGCTCAGCTTCATCTTT 584
      115  LeuProValMetValTyrIleHisGlyGlyLysLeuMetValGlyGlyAlaSerThrTyr 134

```

QY	585	GATGGGTCCGGCTTGGCTGTATATGAGACGTGCTGTGTGTGTGTGTGTATCCAGTACCGGCTA	644
Db	135	AerGlyLeuAlaLeuSerLahIleGluAenValValValThrIleGlnTyArgLeu	154
QY	645	GSATATTTGGTTTCTTACCCACACAGGATGAGCATGTGCTCCGGGAACTGGGCTTCAG	704
Db	155	GlyIleGlyGlyPheGlyPheAenIleAerGlu-----Leu	166
QY	705	GACCAAGTGGCTGCTCTGTCTCGGTCCAGAAAGACATCGATTTCTTCGTGGGAGACCC	764
Db	167	PheLeuValAlaValAenArgTrpValGlnAerAenIleAlaAerPheGlyAerPro	186
QY	765	AGCTGTGACACATTTTGGGAGATCCGGGAGACCCATAAGTTTCTTACTTTATAGT	824
Db	187	GlySerValThrIlePheGlyGlnSerAlaGlyGlnSerValSerIleLeuLeuLeu	206
QY	825	TCTCCCATGGCCAAAGGCTTATTCCAAAGACCATCATGAGAGTGGGGGTGGCCATATC	884
Db	207	SerProLeuThrLysAenLeuPheAenArgAlaIleSerGlnSerGlyAlaAlaLeuLeu	226
QY	885	CCTTACCCGGAGGCCCATGATTATGAAAGAAGTGAGAC-----CTGAGGTGTGCA	938
Db	227	SerSerLeu-----PheArgLysAenThrLysSerLeuAlaGlnLysIleAla	242
QY	939	CATTCTGTGTGTAAACAATGCGTAGATCTGAGAGCCCTGAGGCGGTGAGGACAA	998
Db	243	IleGluAlaGlyCysLysThrTrpThrIleSerAlaValMetValHisCysLeuArgGlnLys	262
QY	999	CCCTCCAAAGAGCGTCGT-----ACCTTCAGCCAGAA-----	1031
Db	263	ThrGlnGlnGluLeuMetGluValThrLeuLysMetLysPheMetAlaLeuAerLeuVal	282
QY	1032	-----ACAAAGTCTTCACTCGAGTGGTGTGATGAGTGTCTTCTTCT	1073
Db	283	GlyAerProLysGlnAenThrAlaPheLeuThrTrpValIleAerGlyValLeuLeuPro	302
QY	1074	AATGACCTCTAGATCTATTGTCTCAAAAAGCATTTAAAGCAATTCCTTCATCATGCGA	1133
Db	303	LysAlaProAlaGlnLleTyGlnGlnLysLysTyAenMetLeuProTyMetValGly	322
QY	1134	GTCAATACACAGAGTGTCCTCCGCGCGCTAG-----	1169
Db	323	IleAenGlnGlnLysPheGlyTrpIleIlePheMetGlnMetLeuGlyTyProLeuSer	342
QY	1170	-----AAGAGGCTCTGAGATCTCACTAGTGGCTCCAAAGTCCCT	1211
Db	343	GlnGlyLysLeuAerPheGlnLysThrAlaThrGlnLeuLeuTrpLysSer-----	358
QY	1212	GCCCTCACTGATACAAACACTCCCTGGACATCCCGCTCAGTATTGGACCTTGGGCT	1271
Db	359	-----TyProIleValAenValSerLysGlnLeuThrProValAlaThr	373
QY	1272	AATGAATCTTCCATGACACAGACCTCCCTGACATAAGATCCAGACAGCTTTCGAACTTG	1331
Db	374	GlnLysTyLysLeuGlyGlyThrAerAerProValLysLysLysAerLeuPheLeuAerMet	393
QY	1332	CTTGGAGATGAGTCTTTGTGTGCTCCGTCACATGATCAACAGTGTATCAACAGATGCT	1391
Db	394	LeuAlaAerPheLeuPheGlyValProSerValAenValAlaArgHisArgAerAla	413
QY	1392	GGTGCACCTGTCTATCTTATGATGTTTCCGACCCGAGTCAAGTCTTGAAGACAGAG	1451
Db	414	GlyAlaProThrTyMetTyGlnTyArgTyArgProSerPheSerGlnAerPheMetArg	433
QY	1452	CCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGTTGTGTGCTGGGTGCTTC	1511
Db	434	ProLysThrValIleGlyAerPheArgLysArgGlnIlePheSerValLeuGlyAlaProPhe	453
QY	1512	CTGAAGGGGAGCATTTGATGTTCAAGAGGACACGAGGAGAGAAAGTTATCGACCCGG	1571
Db	454	LeuLys-----GlnGlyAlaThrGlnGlnGlnLysLysLeuSerLys	467

```

QY      1572  AAGTATGTAATATCTGGGCTTACTTGTCTGAAACCGGGAACTCTTAATGGAAAGACCTG  1631
Db      468  MetValMetIysTyrTrpAlaAsnPheAlaArgAsnGlyAsnProAsnGlyValGlyLeu  487
QY      1632  TCTCTGGCGACGCTTAATATCTGACGTGACGAGTACCTCCGACCTGACTTGAACATGAGC  1691
Db      488  ProGlnTrpProIaIaTyrAspTyrIysGlnGlyTyrLeuGlnIleGlyAlaThrGln  507
QY      1692  CTCGGAAGAGACTTCAAAGAACCGCGGTGATTTTGGACC  1733
Db      508  AlAlaGlnIlyLeuIysAspIysGluValAlaIaPheTrpTr  521

RESULT 13
US-08-446-100-26
; Sequence 26, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glena
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
US-08-446-100-26

Alignment Scores:
Pred. No.: 3 446-83 Length: 454
Score: 942.50 Matches: 200
Percent Similarity: 59.33% Conservative: 64
Best Local Similarity: 44.94% Mismatch: 150
Query Match: 24.34% Indels: 31
DB: 3 Gaps: 6

US-10-023-515-1 (1-2158) x US-08-446-100-26 (1-454)
QY      468  GAGAGTGGCTTACTTGAACATCTAGCGGCTCCAGCCGATACAGGCTTCAAGCTC  527
Db      1  GlnHisCysLeuTyrLeuAsnIleTyrThrProAlaAspLeuThrIlyValAsnArgGln  20

```

[illegible]

```

Db      3712  |||:::|||||  |||:::  ::|||:::|||||  |||:::
Db      1602  CGAACCCGGGAATCTTAATGGAGACGCTGCTGTGGCCAGCTTATATCTGACTGAG 1661
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      392   ArgangniyaenProaenGlyLeuProhiStrProGluTyraenGlnlyeGlu 411
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      1662  CAGNACCTCCGCTGAGCTTGAACATGAGCGCTCGGACGAGAGACTCAAGAACCCCGGCT 1721
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      412   G1yTyLeuGln1leG1yAlaAsnThrGln1Ala1aGlnTybLeuYbAsP1yG1uAl 431
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy      1722  GATTTTGACCAAGC 1736
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      432   Alaphetrinraen 436
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 14
US-08-446-100-28
; Sequence 28, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Bloomfield, Clarence A
; APPLICANT: Willard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenma
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Bloomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
; US-08-446-100-28

Alignment Scores:
Pred. No.: 3,44e-83 Length: 454
Score: 942.50 Matches: 200
Percent Similarity: 59.55% Conservative: 65
Best Local Similarity: 44.94% Mismatches: 149
Query Match: 24.34% Indels: 31
DB: 3 Gaps: 6

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 08:53:20 ; Search time 148.694 Seconds

(without alignments)
11126.674 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 3872
Sequence: 1 ccaagcgcgcgcaaaacagc.....aaaaaaaaaaaaaaaaaaaaa 2158

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 3420798

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlh
-Q=/cgn2_1/USPTO/6664091/runat_14062005_133501_9433/app_query.fasta_1.4238
-DB=Published Applications AA -OFMT=faetan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOCL=0 -LOOEXT=0 -UNITs=bits -START=1 -END=-1 -MATRIX=bloum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=6664091.@CGN_1_1_295 @runat_14062005_133501_9433
-NCPU=6 -ICPU=3 -NO_MMAP -LANG=ODTER -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11B_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	79.5	581	13	US-10-023-515-2
2	3079	79.5	581	15	US-10-674-636-2
3	3079	79.5	581	16	US-10-757-262-46
4	2915	75.3	575	15	US-10-451-168-91
5	2900	74.9	642	15	US-10-433-256-10
6	2607.5	67.3	581	15	US-10-451-168-92
7	2606	67.3	525	15	US-10-094-749-2375
8	2597.5	67.1	581	15	US-10-114-270-196
9	2588.5	66.9	618	15	US-10-381-898-2
10	2454	63.4	462	15	US-10-104-047-2219
11	1985	51.3	549	15	US-10-233-933A-2
12	1985	51.3	542	16	US-10-233-933A-4
13	1780	46.0	356	15	US-10-451-168-93
14	1230	31.8	549	9	US-09-925-298-689
15	1230	31.8	549	14	US-10-102-806-689
16	1230	31.8	550	15	US-10-267-756-5
17	1230	31.8	559	16	US-10-858-271-28
18	1186.5	30.6	554	9	US-09-895-860-4
19	1186.5	30.6	554	15	US-10-377-072-4
20	1186.5	30.6	554	16	US-10-377-072-4
21	1179.5	30.5	571	10	US-09-931-836-23
22	1179.5	30.5	571	11	US-09-833-245-1090
23	1179.5	30.5	571	13	US-10-036-342-23
24	1179.5	30.5	571	13	US-10-036-041-23
25	1179.5	30.5	571	14	US-10-028-072-542
26	1179.5	30.5	571	14	US-10-035-855-23
27	1179.5	30.5	571	14	US-10-140-808-542
28	1179.5	30.5	571	14	US-10-121-049-542
29	1179.5	30.5	571	14	US-10-123-904-542
30	1179.5	30.5	571	14	US-10-140-470-542
31	1179.5	30.5	571	14	US-10-175-746-542
32	1179.5	30.5	571	14	US-10-176-918-542
33	1179.5	30.5	571	14	US-10-176-921-542
34	1179.5	30.5	571	14	US-10-227-884-210
35	1179.5	30.5	571	14	US-10-036-214-23
36	1179.5	30.5	571	14	US-10-137-865-542
37	1179.5	30.5	571	14	US-10-140-474-542
38	1179.5	30.5	571	14	US-10-035-719-23
39	1179.5	30.5	571	14	US-10-142-431-542
40	1179.5	30.5	571	14	US-10-143-114-542
41	1179.5	30.5	571	14	US-10-230-163-210
42	1179.5	30.5	571	14	US-10-036-160-23
43	1179.5	30.5	571	14	US-10-230-338-210
44	1179.5	30.5	571	14	US-10-142-419-542
45	1179.5	30.5	571	14	US-10-218-631-210

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A. J.
; TITLE OF INVENTION: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:
Pred. No.: 3,396-269 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.52% Indels: 0
DB: 13 Gaps: 0

US-10-023-515-1 (1-2158) x US-10-023-515-2 (1-581)

QY 96 ATGCACAGGAGCTTACTGATCTGCTTCAACAAAGTGCTTTTCTGATTTCCAGCCC 155
DB 1 MetProGlnGlyLeuThrSerSerLeuSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
QY 156 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
DB 21 LeuLeuGlnIleSarGlnTrpGlnIleThrGlyProSerAlaGlnGlyProGlnAlaGln 40
QY 216 ACAAGGCTGGAGTGGATTCAGGGGACCAAGTCACTGTGCTGGGAAAGCCCTGTGCTGTG 275
DB 41 ThrArgLeuGlnIleTrpIleGlnGlyLeuValThrValLeuGlnIleSerProValProVal 60
QY 276 AACGTGTTCTCGGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGGGATTAGAAC 335
DB 61 AsnValPheLeuGlnIleProPheAlaAlaProProLeuGlnIleSerLeuArgPheThrAsn 80
QY 336 CCGGACCTGGCATCCGCTGGGATTAATTGCGAAGAACCACTCTACCCCTTAATTGTGC 395
DB 81 ProGlnProIleSerProTrpAsnLeuArgGlnAlaThrSerTrpProAsnLeuGln 100
QY 396 CTCCAGAACTCAGAGTGGCTGCTCTTAAGATCAACAATGCTCAAGGTGATTAACCGAAA 455
DB 101 LeuGlnAsnSerGlnTrpLeuLeuAsnArgGlnIleMetLeuIleValIleTrpProIle 120
QY 456 TTGCGAGTGTCAAGAGACTGCTTACCTGAACATGTATGCGCTGCGCCACCGCATACA 515
DB 121 PheGlnValSerGlnIleAspCysLeuTrpLeuAsnIleTrpAlaProAlaIleValAspThr 140
QY 516 GGCTCAAGCTCCCGCTTGGTGGTGGCTTCCAGAGAGTGGCTTCAAGACTGGCTCACCC 575
DB 141 GlnSerIleLeuProValLeuValIleTrpPheProGlnIleValPheLeuThrGlnIleVal 160
QY 576 TCCATCTTTGATGGTCCGCTGCTGCTGCTTGAAGAGAGTGGCTGTTGGTGGCTGAC 635
DB 161 SerIlePheAspGlnSerAlaLeuAlaIleValIleValIleValIleValIleValIle 180
QY 636 TACGGGCTAGAAATATTTGGTTTCTTCAACACATGGGATCAAGCATGCTCCGGGAACTGG 695
DB 181 TyrArgLeuGlnIlePheGlnIlePhePheThrTrpAspGlnIleValIleProGlnIleVal 200
QY 696 GCGTTCAAGAGACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
DB 201 AlaPheIleAspGlnValIleAlaIleLeuSerTrpValGlnIleValIleValIleValIle 220
QY 756 GGGGACCCCACTGCTGTGACCATCTTTGGGAGTCCGCGGAGGCATTAAGTGTTCAGT 815
DB 221 GlnAspProSerSerValThrIlePheGlnIleSerAlaIleIleSerValSerSer 240
QY 816 CTTATACGTGCTCCCATGAGCCAAAGGCTTATTCACAAAGCATCATGAGAGAGTGGGGTG 875
DB 241 LeuIleLeuSerProMetAlaIleValGlnIleLeuPheIleValIleValIleMetGlnSerGlnIle 260
QY 876 GCCATCATCCCTTAACCTGAGAGCCCATGATTATGAGAAAGTGAAGAGCTGAGAGTGGT 935
DB 261 AlaIleIleProIleTrpLeuGlnAlaIleIleAspTrpGlnIleSerGlnIleValIleVal 280
QY 936 GCACATTTCTGTGTACATAGCTCAGACTCTGAGGCCCTGCTGAGGCTCTGAGGAGCA 995
DB 281 AlaIlePheCysGlnAsnAsnAlaSerAspSerGlnIleLeuLeuArgCysLeuAlaGlnTrp 300
```

```

QY 996 AAACCTCCAGAGAGCTGCTGACCTTCAACCCAGCCGAAAAACAAGTCTTTCACTGAGTGT 1055
DB 301 LysProSerIleGlnLeuLeuThrLeuSerGlnIleValIleTrpIleValIleValIleVal 320
QY 1056 GATGCTGCTTTTCTTCTTAATGAGCCTCAGATCTATGTTGCTCAGAAAGCATTTAAACA 1115
DB 321 AspGlnAlaPhePheProAsnIleProLeuAsnIleLeuSerGlnIleValIleValIleVal 340
QY 1116 ATTCTCCATCATATGGAGTCAATTAACACAGAGTGGCTTCTGCTGCTGCTGCTGCTGCT 1175
DB 341 IleProSerIleIleGlnValIleAsnAsnIleGlnIleValIleValIleValIleValIle 360
QY 1176 GCTCTGAGATCTCAGTGGCTTCCAAAGTCCCTTGGCCCTTCATGATGATCAAAACATC 1235
DB 361 AlaProGlnIleLeuSerGlnIleValIleValIleValIleValIleValIleValIleVal 380
QY 1236 CTGCACATCCCGCTCAGATTTTTCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1295
DB 381 LeuIleIleProProGlnIleValIleValIleValIleValIleValIleValIleValIle 400
QY 1296 TCCCTGACTGAATCCGACAGTCTTCTGAGCTTGTGAGATGTTGTTGTTGTTGTTGTTGTT 1355
DB 401 SerLeuThrGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 420
QY 1356 CCTGCACTGATCAGAGCTCGATATCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
DB 421 ProAlaLeuIleIleThrAlaArgTrpIleValIleValIleValIleValIleValIleVal 440
QY 1416 TTTCGGGACCGGCTCAGTGTCTTGAAGACAGCAAGCGGCTTTTGTCAAGCCGACAC 1475
DB 441 PheArgIleSarProGlnCysPheGlnAspThrIleValIleValIleValIleValIleVal 460
QY 1476 GCTGATGAGTCCGCTTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1535
DB 461 AlaAspGlnIleValIleArgPheValIleGlnIleValIleValIleValIleValIleVal 480
QY 1536 GAAGGACCCAGAGAGAGGAGAACTTACGAGCGCGAAGATGATGAATATAGAGCTGCTGCT 1595
DB 481 GlnIleValIleThrGlnIleGlnIleValIleLeuSerIleValIleValIleValIleValIle 500
QY 1596 TTGCTCGAAGCCGGAATCTTAATGGAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1655
DB 501 PheAlaArgThrGlnIleAsnProAsnGlnIleValIleValIleValIleValIleValIle 520
QY 1656 ACTGAGCAGTACTCTCAAGCTGACCTTGAACATGAGCCTTGGACAGAGCTCAAGAACCG 1715
DB 521 ThrGlnGlnIleThrLeuGlnIleLeuAsnIleMetSerLeuGlnIleValIleValIleVal 540
QY 1716 CCGGTGATTTTGGACAGACAGCAACATCCCGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 1775
DB 541 ArgValAspPheThrPheSerThrIleProIleIleLeuSerIleValIleValIleValIle 560
QY 1776 AGTCTCTTTTCTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1835
DB 561 SerProLeuSerSerLeuThrPheLeuSerIleLeuGlnIleProPhePhePheCysAla 580
QY 1836 CCT 1838
DB 581 Pro 581

RESULT 2
US-10-674-636-2
; Sequence 2, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A. J.
; APPLICANT: Siles-Santiago, Imaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLSTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674, 636
; CURRENT FILING DATE: 2003-09-29
```


PRIOR APPLICATION NUMBER: US/10/023, 515
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,369
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/279,508
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 581
TYPE: PRT
ORGANISM: Homo sapiens
US-10-674-636-2

Alignment Scores:

Pred. No.:	3,39e-269	Length:	581
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	79.52%	Indels:	0
DB:	15	Gaps:	0

US-10-023-515-1 (1-2158) x US-10-674-636-2 (1-581)

96 ATGCCAGAGGACTTACTCATCTGCTTCAACAATGGTGGCTTTTCTGATTCACAGCC 155
1 MetProGlnGlyLeuTherSerSerAlaSerGlnTrpCyPhePheLeuLeuGlnPro 20
156 CTGTTGGGACACAGACAGTGGGAAAAAATGGGCTTCTGCTGAAGAGGCCACAGAGAAC 215
21 LeuLeuGlnAlaArgGlnTrpGlySerThrGlyProSerAlaGlnGlyProGlnArgAsn 40
216 ACCAGGCTGGGATGATTCAGAGGACAGCAAGTCACTGTGCTGGAAGCCCTGCTGCTG 275
41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
276 AACGATGCTCCGAGATCCCTTTGCTGCTCCCGCTGGAGATCCCTGAGATTACGAAC 335
61 AsnValAlaPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
336 CCGGAGCTGCATCGCCCTGGAGATACTTGGAGAAAGCCACTCTTAATTTGTC 395
81 ProGlnProAlaSerProTrpPhePheLeuArgGlnAlaThrSerLysProLeuLeuGly 100
396 CTCGCAAGACTCAGAGTGGCTGCTTCAATCAACATGCTCAAGTGGATTAACCGAA 455
101 LeuGlnAsnSerGlnTrpLeuLeuAspGlnHisMetLeuLysValHisTrpProLys 120
456 TTGGAGTGTCAAGAGACTGCTCTACCTGAACATCTATGCGCTGCGCCAGCCGATACA 515
121 PheGlyValSerGlnAspGlyLeuValTrpPheProGlyAlaPheLysThrGlySerAla 140
516 GGCCTCAAGCTCCCGCTTGGTGTGGTTCAGAGAGTGGCTTCAAGATGGCTCAAGCC 575
141 GlySerLysLeuProValLeuValTrpPheProGlyAlaPheLysThrGlySerAla 160
576 TCCATCTTTGAAGGGTCCCGCTGGCTGCTATGAGAGAGTGTGTGTGTGCTGCTG 635
161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValAlaValGln 180
636 TACCGGCTAGGAATATTGGTTCCTTCAACATGAGGATCAGCATGCTCCGGGAACTGG 695
181 TyrArgLeuGlyIlePheGlyPhePheTrpTrpPheGlnHisAlaProGlyAsnTrp 200
696 GCGTTCAAGAGACAGATGGCTGCTGTGCTGGGCTCAGAGAAACAATCAGATTCTTCGGT 755
201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGlnPhePheGly 220
756 GGGGACCCCAAGCTCTGTGACATCTTGGCAGAGTCCGGGAGCCATAAGTGTTCAT 815
221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
816 CTTATACGTCTCCATGCGCAAGGCTTATTCACAAAGCATCATGAGATGGGGGTG 875

241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGlnSerGlyVal 260
876 GCCATCATCCCTTACTGAGAGCCCATGATTATGAGAAAGAGAGACCTGACGCTGTT 935
261 AlaIleIleProLysLeuGlnAlaHisAspTrpGlyLysSerGlnAspLeuGlnVal 280
936 GCACATTTCTGTGTAAACAATGCGTCAAGACTGAGAGCCCTGCTGAGTGGCTGAGACA 995
281 AlaHisPheCysGlyLysAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
996 AAACCTCCAGAGACTGCTGACCTTCAGCCAGAAACAAAGCTTTTCACTGAGTGGTT 1055
301 LysProSerLysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
1056 GATGCTGCTTTCTTTCTTAATAGAGCTCTAAGATCTATGCTCAGAAACATTTAAAGA 1115
321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
1116 ATTCCCTTCATCATGAGATCAATTAACAACAAGTGGCTTCTGCTGCTGCTGCTGAGAG 1175
341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLysGln 360
1176 GCTCCTGAGATCCTGAGTGGCTCAACAAGTCCCTTGCCTCATCTGATTAACAACATC 1235
361 AlaProGlnIleLeuSerClySerAsnLysSerLeuAlaLeuHisIleuIleGlnAsnIle 380
1236 CTGCACATCCCGCTCAGATTTTGGACCTTGGGCTTAAGATCTCCATGACAGACAC 1295
381 LeuHisIleProProGlnTrpIleuHisLysValAlaAsnGlnLysTrpHisAspLysHis 400
1296 TCCCTGACTGAATCCAGACAGTCTTGTGACTGCTGCTGAGAGATGTTCTTTGGGTC 1355
401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyLysPheValPheValVal 420
1356 CCTGACATGATCAACAAGCTGATATCAACAAGATGCTGTGACCTGTCTACTTATAGAG 1415
421 ProAlaLeuIleThrAlaArgTrpHisAspArgAlaGlyAlaProValTrpPheTrpGln 440
1416 TTTGGCACCGGCTCAGTGGCTTTGAAGACAGAAAGCCGCTTTGTCAAAGCCGACAC 1475
441 PheArgHisAspProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
1476 GCTGATGAAGTCCGCTTGTGTGGTGGTGGCTTCTGGAAGGGGACATTTGTTATGTC 1535
461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyLysPheValMetPhe 480
1536 GAAGAGCCACGAGAGAGAGAGATTAAGTACGACCGAAGATGATGAATATGCGGTACC 1595
481 GlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpAlaThr 500
1596 TTTGCTCGAACCAGGAATCTTAATGGAACAGACTGTCTGTGGCAGCTTATATCTG 1655
501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
1656 ACTGAGAGTACTCTCAGTGAAGCTTGAACATGAGCTTGGACAGAGACTCAAGAACCG 1715
521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
1716 CCGGTGATTTTGGACACAGACATCCCTGATCTGTCTGCTGCTGCTGCTGACATGCTCAC 1775
541 ArgValAspPheTrpPheThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
1776 AGTCTCTTTTCTTCTTAACTTTCTCTCTCTCTCTCTGACCTTTCTTTTCTTTGTGCT 1835
561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
1836 CCT 1838
581 Pro 581
RESULT 3
US-10-757-262-46

```
/ Sequence 46, Application US/10757262
/ Publication No. US20040197825A1
/ GENERAL INFORMATION:
/ APPLICANT: Karichet, Venkateswarlu
/ APPLICANT: Eliasof, Scott D.
/ APPLICANT: Silos-Santlago, Imaculada
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
/ TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
/ TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
/ TITLE OF INVENTION: 35010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
/ TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
/ TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
/ TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
/ TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
/ FILE REFERENCE: MP103-0071R1R0N0N1M
/ CURRENT APPLICATION NUMBER: US/10757,262
/ PRIOR FILING DATE: 2004-01-14
/ PRIOR APPLICATION NUMBER: US 60/440,318
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/444,783
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: US 60/457,901
/ PRIOR FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: US 60/468,775
/ PRIOR FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/471,614
/ PRIOR FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: US 60/478,742
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/488,529
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/491,156
/ PRIOR FILING DATE: 2003-07-30
/ PRIOR APPLICATION NUMBER: US 60/499,594
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 60/506,332
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 46
/ LENGTH: 581
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-757-262-46
Alignment Scores:
Pred. No.: 3,39e-269 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.52% Indels: 0
Gaps: 0
DB: 16
US-10-023-515-1 (1-2158) x US-10-757-262-46 (1-581)
QY 96 ATGCCAGAGGAGTCACTTCACTGCTTCAAGATGCTCTTTCTGATTTCCAGACC 155
DB 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
QY 156 CTGTGGAGACACAGACAGTGGGGAAGAACTGGGCTTCTGCTGAAGGACACAGAGGAA 215
DB 21 LeuLeuGlnIleAlaGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnAlaGln 40
QY 216 ACCAGGCTGGAGATGATTCAGGCAAGCAAGATCATCTGTGCTGGAGACCTGTGCTGTG 275
DB 41 ThrArgLeuGlnGlyTrpIleGlnGlyLeuGlnAlaThrAlaLeuGlnGlySerProValProVal 60
QY 276 AACGTGTTCTCGAGATCCCTTTGCTGCTCCCGCTGGAGATCCCGGCAATTTAAGAAC 335
DB 61 AsnValPheLeuGlnGlyValProPheAlaIleAlaProIleuGlnGlySerLeuAlaGlnPheThrAsn 80
```

```
QY 336 CCGAGGCTTCATCGCCCTGGGATTAACCTTGCAAGAACCCATCCCTAATTTGTGC 395
DB 81 ProGlnProIleAlaSerProIleAlaGlnAlaThrSerLeuProAlaLeuGln 100
QY 396 CTCAGAACTCAGAGGCGGCTTGAATCAACATCTTAAGAGTGCATTTACCGGAA 455
DB 101 LeuGlnAlaSerGlnTrpLeuLeuLeuSerGlnIleMetLeuValAlaIleTrpAla 120
QY 456 TTGGAGTGTCAAGAACTGCTTCACTGAACATCTATGCGCGCTCCAGCGCATACA 515
DB 121 PheGlyValSerGlnAlaSerCysLeuTrpLeuAlaIleTrpAlaProAlaIleAlaSer 140
QY 516 GGTCTCAAGCTCCCGCTTGGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGCTCAGCC 575
DB 141 GlySerLeuLeuProValLeuValAlaTrpPheProGlnGlyAlaPheLeuThrGlySerAla 160
QY 576 TCCATCTTTGATGGTCCGCCCTGGCTGCTCATGAGAGCTGCTGTTGTGTCAG 635
DB 161 SerIlePheAlaSerGlySerAlaLeuAlaIleTrpGlnAlaSerValLeuValAlaGln 180
QY 636 TACCGGCTAGGAATATTTGTTTCTTCAACCATGGGATCAGATGCTCCGGGAACTGG 695
DB 181 TyrArgLeuGlnGlyIlePheGlyPhePheThrTrpAlaSerGlnIleAlaProGlyAlaSer 200
QY 696 GCTTCAAGAACACAGTGGCTGCTGTCTGCTGGGTCCAGAAACATCGAGTTCCTCGGT 755
DB 201 AlaPheLeuAlaSerGlnValAlaIleAlaLeuSerTrpValGlnIleAlaSerValPheGly 220
QY 756 GGGGACCCGAGCTGTGAGACCATTTTGGAGAGTCCGGGAGCCATAGTGTGTTAGT 815
DB 221 GlnAlaProSerSerValAlaIlePheGlnGlnSerAlaGlyAlaIleSerValSerSer 240
QY 816 CTATAGTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGG 875
DB 241 LeuIleLeuSerProMetAlaIleGlyLeuPheAlaIleAlaLeuGlnSerGlyAla 260
QY 876 GGCATCATCCCTTACCTGAGAGGCCCATGATTATGAGAAAGTGAAGCACTGAGTGT 935
DB 261 AlaIleIleProTrpLeuGlnAlaIleAlaSerGlyGlnIleSerGlnAlaPheLeuGlnVal 280
QY 936 GCATATTTCTGTGTAAACAATGCGTCAAGCTCGAGGCGCTGAGGCTGAGAGCA 995
DB 281 AlaIlePheCysGlyAlaSerAlaSerAlaSerAlaLeuAlaLeuAlaGlyValAlaIle 300
QY 996 AAACCTCCCAAGAGCTGCTGAACCTCAGCCAGAAACAAAGCTTTTCACTGAGTGT 1055
DB 301 LysProSerLeuGlnIleLeuLeuThrLeuSerGlnIleTrpLysSerPheThrArgValAla 320
QY 1056 GATGATGCTTTCTTTCTTAATAGGCTCAGATCTATGCTCAGAAAGCATTTAAACA 1115
DB 321 ArgGlyAlaIlePhePheProAlaSerGlnIleProLeuAlaSerLeuSerGlnIleAlaPheLeuAla 340
QY 1116 ATTCTTCCATATGAGAGTCAATTAACACAGAGTGTGCTTCTGCTGCTCATGAGAG 1175
DB 341 IleProSerIleIleGlyValAlaSerAlaIleGlnIleAlaSerGlnIlePheLeuLeuProMetLeuGln 360
QY 1176 GCTCCGAGATCTAGAGGCTCCCAAGATCCCTGCGCTCCATCTGATTAACAAATC 1235
DB 361 AlaProGlnIleLeuSerGlnIleSerAlaIleSerLeuAlaLeuIleGlnAlaIle 380
QY 1236 CTGCAATCCGCGCTCAGATATTTGACCTTGTGCTAATGAATCTTCCATGACAC 1295
DB 381 LeuIleIleIleProGlnIleTrpLeuIleAlaValAlaSerGlnIleTrpPheIleAlaSerVal 400
QY 1296 TCCCTGATGAAATCCAGAGACAGATCTTCTGAGCTTGTGAGATGATCTTTGTGTC 1355
DB 401 SerLeuTrpGlnIleAlaGlnSerSerLeuLeuAlaSerLeuGlnIleAlaPhePheValVal 420
QY 1356 CTTGCACTGATCAGAGCTGATATCAAGAGATGCTGCTGACCTGATCTTATGAG 1415
DB 421 ProAlaLeuIleThrAlaIleAlaGlyTrpAlaIleAlaSerAlaIleAlaProValIleTrpPheTrpGln 440
QY 1416 TTTGGGACCGGCTCAGTGCCTTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACAC 1475
```

```
|||||
Db 441 PheArgHisArgProGlnCysPheGluAaPThrLysProAlaPheValLysAlaAspHis 460
Qy 1476 GCTGATGAAGTCCGCTTTTGTTGCTGCTGCTTCTGAAGGGGAGCAATGTTAATGTTTC 1535
Db 461 ALaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
Qy 1536 GAAGGAGCCAGGAGGAGGAGAACTAGAGCCGAGAGATGATAAATAGTGGGATACC 1595
Db 481 GluGlyAlaThrGluGluGluLysLeuLeuSerAlaGlyMetMetLysTyrTrpAlaThr 500
Qy 1596 TTGTGCGAAGCCGGGAAATCTTAATGGAAAGCACTTCTGTGGCCAGCTTAATAATCTG 1655
Db 501 PheAlaArgThrGlyAsnProAlaGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeu 520
Qy 1656 ACTGAGCACTACTCCAGCTGAGCTTGAACATGAGCTTGAGACAGACTCAAGAACCG 1715
Db 521 ThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Qy 1716 CGGGTGGATTTTGGACGACCAATCCCGCTGATCTGTGCTGCGACATGCTCCAC 1775
Db 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
Qy 1776 AGTCCCTTTCTTCCCTTAATCTTCTCTCTCTCCAGCTTTCTTTTCTTTGTGCT 1835
Db 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
Qy 1836 CCT 1838
Db 581 Pro 581

RESULT 4
US-10-451-168-91
/ Sequence 91, Application US/10451168
/ Publication No. US20040091969A1
/ GENERAL INFORMATION:
/ APPLICANT: SMITHKLINE BEECHAM CORPORATION
/ APPLICANT: SMITHKLINE BEECHAM P.L.C.
/ TITLE OF INVENTION: NOVEL COMPOUNDS
/ FILE REFERENCE: GP50039
/ CURRENT APPLICATION NUMBER: US/10/451,168
/ CURRENT FILING DATE: 2003-11-12
/ PRIOR APPLICATION NUMBER: PCT/US01/49232
/ PRIOR FILING DATE: 2000-12-17
/ PRIOR APPLICATION NUMBER: 60/256,710
/ PRIOR FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 60/257,048
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 60/260,482
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: 60/264,922
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/266,797
/ PRIOR FILING DATE: 2001-02-06
/ PRIOR APPLICATION NUMBER: 60/276,988
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/281,535
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/289,622
/ PRIOR FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 91
/ LENGTH: 575
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-451-168-91

Alignment Scores:
Pred. No.: 2,37e-254 Length: 575
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservative: 3
```

```
Best Local Similarity: 97.02% Mismatches: 6
Query Match: 75.28% Indels: 8
DB: 15 Gaps: 1
US-10-023-515-1 (1-2158) x US-10-451-168-91 (1-575)
Qy 129 TGGTGTCTTTCTGATATCTCCAGCCCTGTGGGACACAGACAGTGGGAAAACTGGG 188
Db 14 TrpAlaIleTrpValLeuAlaAlaPro-----ThrLysGly 25
Qy 189 CCTTGTCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGCAAGTC 248
Db 26 ProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrIleGlnLysGlnVal 45
Qy 249 ACTGTGCTGGGAAGCCCTGTGCTGGAACGAGTTCCTCGAGAGTCCCTTGTGCTGCC 308
Db 46 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 65
Qy 309 CCGCTGGATCCCTCGATTTACGAACCCGACGCTGCATCCGCTGGGATTAATTCGGA 368
Db 66 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 85
Qy 369 GAAGCCACTCTTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTAATCAA 428
Db 86 GluAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGluTrpLeuLeuAspGln 105
Qy 429 CACATGCTCAAGGTCAATTAACTCCGAAATTCGGAGTGCAGAGCTCCTTAACCTGAAC 488
Db 106 HisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeuAsn 125
Qy 489 ATCTATGCGCTGCCCAAGCCAGCCGATACAGGCTCCAGGCTCCCGCTTGGTGTCCCA 548
Db 126 IleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhePro 145
Qy 549 GGAAGTGCCTTCAAGACTGAGCTCAAGCTTCATCTTGAATGGATCCGCTGCTGCTAT 608
Db 146 GlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIaTyr 165
Qy 609 GAGGACGTGCTGTGTGTGCTCCAGTACCGGCTAGGAATTTGGTTCTTCCACCA 668
Db 166 GluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 185
Qy 669 TGGGATCAGCATGCTCCGGGGAATGGGCTTCAAGACACAGAGTGGCTGCTGCTGG 728
Db 186 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrp 205
Qy 729 GTCCAGAAAGACATGAGTTCTTGGTGGGAGCCCAAGCTGTGACCATCTTGGCGAG 788
Db 206 ValGlnLysAsnIleGluPhePheGlyAspProSerSerValThrIlePheGlyGlu 225
Qy 789 TCCGGGGAGCCATTAAGTGTCTTAAGTCTTAATACGTCTCCATGGCCAAAGCTTAATTC 848
Db 226 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPhe 245
Qy 849 CACAAAGCATCATGAGAGAGTGGGAGGATCATCCCTTACCTGGAGGCCCATGATTAT 908
Db 246 HisLysAlaIleMetGlnSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyr 265
Qy 909 GAGAAGAGTGAAGACCTGAGGTGTTGACATTTCTGTGTAAACATGCGTCAAGCTCT 968
Db 266 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 285
Qy 969 GAGGCCCTGCTGAGGTGCTGAGGACAAACCTTCCAAAGAGCTGTGACCTCAGCCAG 1028
Db 286 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGln 305
Qy 1029 AAAACAAAGCTTTTACCTCGAGTGTGATGATGCTTTCTTCTTAATGAGGCTCTAGAT 1088
Db 306 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 325
Qy 1089 CTATTGTCTCAGAAAGCAATTAAAGCAATTCCTTCATCATGAGTCAATTAACAGCAG 1148
Db 326 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValHisAsnHisGln 345
```

QY 1149 TGTGCTTCTGCTGCTATGAGAGAGCTCCGAGATCTCAAGTCTCAAGCTCC 1208
DB 346 CysGlyPheLeuLeuProMetLeuSglValaProGluIleLeuSerGlySerAsnLysSer 365
QY 1209 CTTCGCTTCATCTGATACAAACATCTGACATCCGCTCACTATTTGCACTCTGTG 1268
DB 366 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnItyLeuHisIleVal 385
QY 1269 GCTAAATGATATCTTCATGACAAAGACCTCCGATGAAATCCGAGACATCTTGAC 1328
DB 366 AlaLeuGluItyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAsp 405
QY 1329 TTGCTTGAAGATGTCTTTGTGTGTGCTCCCTGACATGATCAGCTCGATACAGAGAT 1388
DB 406 LeuLeuGluAspValaPhePheValaIleProAlaLeuIleThrAlaGlyrHisAlaGlyAsp 425
QY 1389 GCTGATGACCTGCTTACTTCTTATGAGATTTGGACACCGGCTCACTGCTTGAAGACAG 1448
DB 426 AlaGlyAlaProValaItyrPheItyrGluPheArgHisArgProGlnCysPheGluAspThr 445
QY 1449 AAGCCGCTTTTGTCAAGCCGACACGCTGATGAAAGCTGCTTGTGTGCTGCTGCC 1508
DB 446 LysProAlaPheValaIlyValaAspHisAlaAspGluValaItyrPheValaIlePheGlyGlyAla 465
QY 1509 TTCCTGAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGAGAAATTACTGAGC 1568
DB 466 PheLeuLysGlyAspIleValaIleMetPheGluGluIlyAlaThrGluIleGluIlyLeuLeuSer 485
QY 1569 CGAAGATGATGAATACTGGGCTACTTGTGCTGAAACCGGAAATCTTAATGGAAACGAC 1628
DB 486 ArgLysMetMetLysItyrIlePheAlaThrPheAlaArgThrIleIyAsnProAsnGlyAsnAsp 505
QY 1629 CTGTCTTGTGGCCAGCTTTTAATCTGACTAGACAGTACTCCAGCTGAGACTTGAACATG 1688
DB 506 LeuSerLeuItyrProAlaItyrAsnLeuThrGluGlnItyrLeuGlnIleuSerLeuAsnMet 525
QY 1689 AGCCTCGAGACAGACTCAAGAAACCGCGGTGAGATTTTGGACCGACCATCCCTCTG 1748
DB 526 SerLeuGlnItyrArgLeuLysGluProGlyValaAspPheItyrHisSerThrIleProLeu 545
QY 1749 ATCTGCTTGTGCTCCGACATGCTCCACAGTCTCTTCTTCTTAACTTCTCTCTCTGC 1808
DB 546 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 565
QY 1809 CTCGAGCTTCTTCTTCTTCTTCTTGTGCTCT 1838
DB 566 LeuGlnItyrPhePhePheCysAlaPro 575

RESULT 5
US-10-433-256-10
Sequence 10, Application US/10433256
Publication No. US20040081960A1
GENERAL INFORMATION:
APPLICANT: SANJANMALA, Madhusudan M.; YAO, Monique G.
APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
APPLICANT: ARVIZU, Chandra S.; RING, HuiJun Z.
APPLICANT: LEE, Ernestine A.; DING, Li
APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
APPLICANT: YUE, Henry; TRIBOUTLEY, Catherine M.
APPLICANT: LU, Dzung Aina M.; LAU, Preeti G.
APPLICANT: WARREN, Bridget A.; YANG, Junning
APPLICANT: CHAKRA, Nandinder K.; NGUYEN, Daniel B.
APPLICANT: GANDHI, Ameena R.; LU, Yan
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-0313 USN
CURRENT APPLICATION NUMBER: US/10/433,256
CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: PCT/US01/47429
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: US 60/254,308
PRIOR FILING DATE: 2000-12-08

Alignment Scores:
Pred. No.: 5,69e-253 Length: 642
Score: 2900.00 Matches: 548
Percent Similarity: 99.82% Conservative: 2
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 74.90% Indels: 0
DB: 15 Gaps: 0

US-10-023-515-1 (1-2158) x US-10-433-256-10 (1-642)

QY 186 GGGCTTCTGCTGAAGGGCCACAGAGAACACGAGCTGGAGATGATTCAGGGACGAA 245
DB 92 GlyProSerAlaGluGluItyrProGlnArgAsnThrArgLeuItyrIleGlnGlyLysGln 111
QY 246 GTCACTGTCTGGAGACCTCTGTGCTGTGAAAGTGTCTCGAGATCCCTTGTGCTGT 305
DB 112 ValThrValLeuGlySerProValaProValaAsnValaPheLeuGlyValaProPheAlaAla 131
QY 306 CCCCCTGGGATCCCTGGGATTTAGAACCCGACCTCGATCCGCTGGAGATTAACCTTG 365
DB 132 ProProLeuGlySerLeuAspGlnPheThrAsnProGlnProAlaSerProIleAspAsnLeu 151
QY 366 CGAAGACCACTCTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTATGAT 425
DB 152 ArgGluAlaThrSerItyrProAsnLeuGlyLeuGlnAsnSerGlnItyrLeuLeuAsp 171
QY 426 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTTACCTG 485
DB 172 GlnHisMetLeuLysValaHisItyrProLysPheGlyValaSerGluAspCysLeuItyrLeu 191
QY 486 AACATCTATGCGCTGCCACCGCCGATACAGAGCTCCAGAGCTCCGCTGCTGCTGCTTC 545
DB 192 AsnIleTyAlaProAlaHisAlaAspThrGlySerItyrLeuProValaLeuValaItyrPhe 211
QY 546 CCAAGAGGCTCTTCAAGCTGCTCAGCTCCATCTTGTATGAGTCCGCTGCTGCTGCC 605
DB 212 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 231
QY 606 TATGAGACGCTGCTGTGTGTGCTGCTCCAGTACCGGCTGAGAAATTTGCTTCTTCAAC 665
DB 232 TyrGluAspValaLeuValaValaValaGlnItyrArgLeuGlyIlePheGlyPhePheThr 251
QY 666 ACATGAGATCAGCATGCTCCGGGAACTGGGCTTCAAGACACAGTGGCTGCTGCTGCC 725
DB 252 ThrItyrAspGlnHisAlaProGlyAsnItyrAlaPheLysAspGlnValaAlaAlaLeuSer 271
QY 726 TGGCTCGAAGAAACATGAGATTTCTGTGGTGGGAGACCCGACTTGTGACCATTTTGGC 785
DB 272 ItyrValaGlnLysAsnIleGlnLysPheGlyGlyAspProSerSerValaItyrIlePheGly 291
QY 786 GAGTCCGGGAGGACATAGTGTCTTGAATCTTATGCTGTCTCCATGGGCAAGGCTTA 845
DB 292 GluSerAlaGlyAlaIleSerValSerSerLeuLeuSerProMetAlaLysGlyLeu 311

846 TTCCAAAGCCATGAGAGAGGAGGAGCCATCATCCCTTAACCTGAGGCCCATGAT 905
Db phenileuylalilemeclusergylvalalalleleprotylleuglulahlaser 331
906 TATGGAAGAGTGAAGACCTGCAAGGTGGTGGACATTTTGTGGTAAACATCGCTGAC 965
Db 332 TGTGTGlysserGlnabPheuglnvalValalahlserGlyabnabnlasaser 351
966 TCTGAGGCCCTGCTGAGGAGCTGAGCAAAACCTCCAAAGAGGTGTGACCCCTGAC 1025
Db 352 SerGlnalaleuueuargyueuuhgthlyrProserLygIueuueuthrueuser 371
906 CAGAAACAAAGCTTTTCACTCGAGTGTGTGATGTGTCTTTCTTAATGAGCCTCTA 1085
Db 372 GlnlystrlyrserPhehrhargValValabgIyAlaPhePheProabngIurProleu 391
1086 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATGTGAGACTCAATAAC 1145
Db 392 AspleuueuserGlnlyserAlaPheLyaballeleProserllelleGlyValabnabnls 411
1146 GAGTGTGGCTTCCTGCTGCTTGAAGAGAGGCTCCGAGATCTCGAGTCCCAACAG 1205
Db 412 GlnCyabgIyPheleuueuPromecluserGlnalProgluValleuSerGlyserabnly 431
1206 TCCCTTGGCCCTCATCTGTATCAAAACATCTGCAATCCGCTCAGTATTGGACCTT 1265
Db 432 SerleuualaleuhsleuileglnabnilleuhsilleProProglInlyrleuhsleu 451
1266 GTGGCTAATGAATCTTCATGACAGACACTCCCTGACTGAATTCGAGACAGTCTTCTG 1325
Db 452 ValalabngIlyrPhehlsabryhslerleuThngIulIargyabserleuueu 471
1326 GACTTGTCTGAGAGTGTCTTGTGTGCTGCTGCTGAGTACAGTCCGATATGACAGA 1385
Db 472 AspleuueuGlyabPvalPhePheValalProalaleuIlethralargytrhsarg 491
1386 GATGCTGTGACACCTGTACTTCTATGAGTTTCGCAACCGGCTCAGTATTGAAAGAC 1445
Db 492 AsphalagIyAlaProvallyrPheTyrgIurPheahghlsargProglInCyserhegluab 511
1446 ACGAAGCCGGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGTGGTGT 1505
Db 512 ThrlyrProalPheVallyleValabPhehsalabargIuValIargPheValPheGly 531
1506 GCTTCTGCAAGGGGAGCAATTGTATGTTGCAAGAGCCAGCAAGAGAGAACTTACG 1565
Db 532 AlaPheleuLyabIyabIleValmeclPheglulalathrgIugIullyserleuueu 551
1566 AGCCGGAAGATGATAAATTAATGAGGCTACCTTGTGCAACCGGGAATCTTAATGGGAA 1625
Db 552 SerahgIyabmecluserlyrTtrpalathrPheahlsargthnGlyabnProabnIyabn 571
1626 GACCTGTCTGTGTGAGCAAGCTTATAATCTGACTGAGCAAGTACTCCAGCTGAGCTTGA 1685
Db 572 AspleuPheleuTrpProalIyabnleuThngIugInlyrleuGlnleuabPheleuabn 591
1686 ATGAGCTTCGACAGAGACTCAAAACCGGGGTGATTTTGGACCAAGCAATCCGCC 1745
Db 592 MetserleuugIyGlnarGlyleuLyabIuProargValaGluPheTrpThserThrllePro 611
1746 CTGATCTGTGTGCTGCAAGCTTCAAGTCCCTTTCTTAACTTAACTTCTCTCT 1805
Db 612 LeuIleuSerAlaserAbPheclleuhslerProleuSerSerleuThrPheleuSer 631
1806 CTCCTCAGCCTTTCTTTTCTTTTGTGTCTCT 1838
Db 632 LeuIleuGlnProPhePhePheCysAlaPro 642

RESULT 6
US-10-451-168-92
; Sequence 92, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:

APPLICANT: SMITHKLINE BEECHAM CORPORATION
APPLICANT: SMITHKLINE BEECHAM P.L.C.
APPLICANT: GLAXO GROUP LIMITED
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP50039
CURRENT APPLICATION NUMBER: US/10/451,168
CURRENT FILING DATE: 2003-11-12
PRIOR APPLICATION NUMBER: PCT/US01/49232
PRIOR FILING DATE: 2000-12-17
PRIOR APPLICATION NUMBER: 60/256,710
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/257,048
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 60/260,482
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 60/264,922
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/266,797
PRIOR FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 60/276,988
PRIOR FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 60/281,535
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/289,622
PRIOR FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 92
LENGTH: 581
TYPE: PRT
ORGANISM: Homo sapiens
US-10-451-168-92
Alignment Scores:
Pred. No.: 1.64e-226 Length: 581
Score: 2607.50 Matches: 499
Percent Similarity: 92.67% Conservative: 7
Best Local Similarity: 91.39% Mismatches: 25
Query Match: 67.34% Indels: 15
Gaps: 3
US-10-023-515-1 (1-2158) x US-10-451-168-92 (1-581)
129 TGGTCTTTTCTCATCTTCACGCCGTGTGGACACAGACAGTGGGAAAACTGGG 188
Db 14 TrrpalalletrPvalleuAlaPro-----ThlyrGly 25
906 CTTTCTGCAAGGGGCAAGAGAAACAGAGCTGGAGTGTGAGGCAAGCAAGTC 248
Db 26 ProserAlagIugIyProglInabnThrargleuGlyTrrpleglnGlyserGlnVal 45
249 ACTGTGTGGAAAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCCCTTGTGCTGCC 308
Db 46 ThrValleuGlyserProvalProValabnValPheleuGlyValProPheAlaPro 65
309 CCGGTGGATCCCTGCAATTAACGAACCGGAGCTGTGAGCCCTGGAGTAACTTGGGA 368
Db 66 ProleuGlyserleuargPheThrsanProglInProalaserProtrParsanleuarg 85
369 GAAGCAACCTCTAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTAATGACA 428
Db 86 GlnalathrserlyrProabnleuCyserGlnabnserGlntrPleuIleuueuabn 105
429 CACATGCTCAAGTGCATTAACCCGAATTTGGAAGTGTCAAAAGACTGCTCTAAGTGA 488
Db 106 HlsmeclleuLyabnIleTyrrProLyabPheGlyAlaserGlnabPheCysleuTytrleuabn 125
489 ATGTATGGGCTCCGCAAGCCGAGTATCAGAGCTCCAAAGCTCCCGTGTGTGGTGGTCCA 548
Db 126 IletyralaProalahlalabserThrslyserlyuabPheProvalleuValItrPhePro 145
549 GGAGGTGCTTCAAGACTGAGTGCAGCTTCATCTTTGATGGGTCCGCCCTGCTGCTAT 608

```

Db      146 GlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIaTyr 165
QY      609 GAGGAGTGTGTTGGTTCCTCCAGTACCGGCTTAGAATAATTTGGTTTCTTCCACACA 668
Db      166 GluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 185
QY      669 TGGGATCAAGCATGCTCCGGGAACTGGGGCTTCAAGGACCAAGGCTGCTGTCTCTGG 728
Db      186 TPrAspGlnHlaIaIaProGlyAsnThrAlaPheLeuAspGlnAlaIaIaLeuSerTyr 205
QY      729 GTCCGAAAGAACATGAGTCTTCGGTGGGACCCCAAGCTGTGACCATCTTTGGCGAG 788
Db      206 ValGlnLysAsnIleGlyPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 225
QY      789 TCCCGGAGGACCATTAAGTGTTCCTAGTCTTACTGTCTCCCAATGGCCAAAGGCTTATTC 848
Db      226 SerAlaGlyAlaIleSerValSerSerLeuLeuSerProMetAlaLysGlyLeuPhe 245
QY      849 CACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGCCCATGATTAT 908
Db      246 HisLysAlaIleMetGlySerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyr 265
QY      909 GAGAAAGTGAAGACCTGAGAGTGTTCACATTTCTGTGGTAAACAATGCGTCAAGCTCT 968
Db      266 GluLysSerGlyAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 285
QY      969 GAGGCTCTGAGTGCCTGAGACAAACCCCTCAAGAGGCTGTGACCTCCACGCCAG 1028
Db      286 GluAlaLeuLeuAspGlyCysLeuAspGlyThrLysProSerTyrSerLeuLeuThrLeuSerGln 305
QY      1029 AAAACAAGCTTTCACTCGATGCTTGAATGCTGCTTTCTTTCTTAATAGCTCTAGAT 1088
Db      306 LysThrLysSerPheThrAlaGlyValAlaAspGlyAlaPhePheProAsnGlnProLeuAsp 325
QY      1089 CATATGCTCAGAAAGCATTTAAAGCAATCTCTCCATCATCGAGTCAATAACCAAGAG 1148
Db      326 LeuLeuSerGlnLysAlaPheLeuValAlaIleProSerIleIleGlyValAsnAsnHisGln 345
QY      1149 TGGGGCTCTGCTGCTTGAAGAGG-----GCTCTGAGATCTCTGAGGCTCCAC 1202
Db      346 CysGlyPheLeuLeuProMetValArgGlyLeuAlaValHisThrAlaThrProSerAsn 365
QY      1203 AAGTCCCTTGCCTC-----CATCTGATCAAAACATCTCGACATCTCCG 1247
Db      366 ArgAspAlaAlaLeuAlaSerThrAlaGlyHisPheHisAspArgHisGlnHisIlePro 385
QY      1248 CCTCAATTTTGCACCTTGTGCTAATGAATCTTCAATGACACAGACCTCTGACTGAA 1307
Db      386 ProGlnTyrLeuHisValAlaAsnGlnTyrPheHisAspLysLeuLeuThrGlu 405
QY      1308 ATCCGACAGCTCTTGCATCTTGGAGATGCTTGTGTGTGCTCCCTGCACTGATC 1367
Db      406 IleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIle 425
QY      1368 AAGCTCGATATCAGAGATGCTGTGCACTGTCTACTTCTTAAGTTCGGACACCG 1427
Db      426 ThrAlaAspGlyHisAspAlaGlyAlaProValTyrPheTyrGlnPheAspHisIleArg 445
QY      1428 CTTGAGTGTGAAGACAGACCGGCTTTTGTCAAGCCGACCAAGCTGATGAAGTC 1487
Db      446 ProGlnCysPheGlnAspThrLysProAlaPheValLysValAspHisAlaAspGlnVal 465
QY      1488 CGCTTTGTGTGCTGTGCTCTCTGAAAGGGGACATGTTATGTTGCAAGAGCCAG 1547
Db      466 ArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlnGlyAlaThr 485
QY      1548 GAGGAGGAGAGTTAATGAGCCGGAAGATGATGAATATCTGGGCTCTTGTGCTGAACC 1607
Db      486 GluGlnGlnLysLeuLeuSerArgLysMetCysTyrTyrAlaThrPheAlaAspGlnHis 505
QY      1608 GGAATCTTAATGGAAGCACTGTCTGTGCGACATTAATCTGAATGAGCACTAC 1667
Db      506 GlyAsnProAsnGlnLysAsnAspLeuSerLeuThrProAlaTyrAsnLeuThrGlnGlnTyr 525

```

```

QY      1668 CTCGAGCTGACCTTGAACATGAGCTCTGGACAGAGACTCAAGAAACCGCGGTGATTTT 1727
Db      526 LeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGlnProAlaArgAspVal 545
QY      1728 TGGACGACGACCATCCCC 1745
Db      546 TrpValThrGlyTyrPro 551

RESULT 7
US-10-094-749-2375
/ Sequence 2375, Application US/10094749
/ Publication No. US20030219741A1
/ GENERAL INFORMATION:
/ APPLICANT: ISOGAI, TAKAO
/ APPLICANT: SUGIYAMA, TOMOYASU
/ APPLICANT: OTSUKI, TETSUJI
/ APPLICANT: WAKAMATSU, AI
/ APPLICANT: SATO, HIROYUKI
/ APPLICANT: ISHII, SHIZUKO
/ APPLICANT: YAMAMOTO, JUN-ICHI
/ APPLICANT: ISONO, YUUKO
/ APPLICANT: HIO, YURI
/ APPLICANT: OTSUKA, KAORU
/ APPLICANT: MAGAI, KENICHI
/ APPLICANT: IRIE, RYOTARO
/ APPLICANT: TAMECHIKA, ICHIRO
/ APPLICANT: SEKI, NAOHICO
/ APPLICANT: YOSHIKAWA, TSUTOMU
/ APPLICANT: OTSUKA, MOTOMYUKI
/ APPLICANT: NAGAHARI, KENJI
/ APPLICANT: MASUHO, YASUHIKO
/ TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
/ FILE REFERENCE: 084335/0160
/ CURRENT APPLICATION NUMBER: US/10/094,749
/ CURRENT FILING DATE: 2002-03-12
/ PRIOR APPLICATION NUMBER: 60/350,435
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: JP 2001-328381
/ PRIOR FILING DATE: 2001-09-14
/ NUMBER OF SEQ ID NOS: 3381
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2375
/ LENGTH: 525
/ TYPE: PRN
/ ORGANISM: Homo sapiens
US-10-094-749-2375

Alignment Scores:
Pred. No.: 2,14e-226
Score: 2606.00
Percent Similarity: 88.77%
Best Local Similarity: 88.25%
Query Match: 67.30%
DB: 15 Gaps: 2

US-10-023-515-1 (1-2158) x US-10-094-749-2375 (1-525)
QY      129 TGGTGTCTTTTCTGATTTCTCCAGCCCTGTGGACACAGACAGTGGGAAAACTGGG 188
Db      14 TrpAlaIleTrpValLeuAlaIaPro-----ThrLysGly 25
QY      189 CCTTGTGCTGAAGGCGCACAGAGAACACAGAGCTGGAGATGGAATGAGGCGAAGCAATC 248
Db      26 ProSerAlaGlnGlyLysProGlnAlaArgAsnThrArgLeuGlyTyrIleGlnGlyLysGlnVal 45
QY      249 ACTGTGCTGGGAAGCCCTGTGCTCTGTGAACGTGTTCCTGGAAGTCCCTTGTGCTGCTCCC 308
Db      46 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIaPro 65
QY      309 CCGTGGGATCCCTGCGAATTTAAGAACCCGACGCTGCACTCGGCTGGGATTAATTGGCA 368
Db      66 ProLeuGlySerIleuAlaGpHeThrAsnProGlnProAlaSerProThrAsnLeuAsp 85

```

QY 369 GAAGCACTCCACCTAATTGTCCTCAGAACTCAGAGTGGCTTATGATCA 428
 DB 86 GUAUATHSerTrProbsnLeuCySeuGlnAsnSerGluTrProLeuLeuAspGln 105
 QY 429 CACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAAAGATGCTCTTACCTGAAC 488
 DB 106 HSmecLeuLysValHisTrProLysPheGlyValSerGluAspYsLeuTrLysLeuAsn 125
 QY 489 ATCTATGCCCTCCGCCACCGCATACAGGCTCCAGACTCCCGCTCTTGGTGTGGTCCCA 548
 DB 126 ILeuTrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhePro 145
 QY 549 GGAAGGTGCTTCAGAACTGAGCTCAGGCTCATCTTTGATGGGCGCGCCGCTGGCTGCTAT 608
 DB 146 GLyGlyAlaPheLysTrnGlySerAlaSerLLeuPheAspGlySerAlaLeuAlaAlaTrp 165
 QY 609 GAGACGTGCTGTGTGTGTCGTCCAGTACCGGCTAGAAATTTGGTTTCTTACACACA 668
 DB 166 GUAAPValLeuValValValGlnTrpArgLeuGlyLLeuPheGlyPhePheThrThr 185
 QY 669 TGGGATCAGCAATGCTCCGGGAACTGGGCTTCAGAGACAGAGTGGCTGCTGTCTGG 728
 DB 186 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrp 205
 QY 729 GTCCAGAAAGAAATGAGTCTTCCGTTGGGGAGCCCGAGCTGTGACATCTTTGGCGAG 788
 DB 206 ValGlnLysAsnLLeuLysPheGlyGlyAspProSerSerValThrLLeuPheGlyGln 225
 QY 789 TCCGGCGGAGCCATAGTGTCTTCTATGTTCTTATCTGTCTCCATGGCCAAAGGCTTATTC 848
 DB 226 SerAlaGlyAlaLLeuSerAlaSerSerLeuLLeuSerProMetAlaLysGlyLeuPhe 245
 QY 849 CACAAAGCCATATGAGAGTGGGGTGGCCATCATCTTACTGGAGGCCATGATTAT 908
 DB 246 HisLysAlaLLeuMetGlySerGlyValAlaLLeuProLysLeuLysAlaHisAspTrp 265
 QY 909 GAGAAAGAGTGAAGACCTCGAGGCTGTGACATTTCTGTGTAAACAATCGTCAAGCTCT 968
 DB 266 GlnLysSerGlnAspLeuGlnValValAlaHisPheCyGlyAsnAsnAlaSerAspSer 285
 QY 969 GAGGCTCTGCTGAGTGTGCTGAGACAAACCTCCAGAGAGCTGTGAACCTCAGCCAG 1028
 DB 286 GUAAlaLeuLeuArgCysLeuArgThrLysProSerLysGlnLeuLeuThrLeuSerGln 305
 QY 1029 AAAACAAAGTCTTTCACTGAGTGGTGTATGGTCTTTCTTCTATGAGCTCTAGAT 1088
 DB 306 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 325
 QY 1089 CTAATGTCTCAGAAAGCAATTTAAAGCAATTCCTCCATCATGAGTCAATAACCAACGAG 1148
 DB 326 LeuLeuSerGlnLysAlaPheLysAlaLLeuProSerLLeuGlyValaAsnAsnHisGln 345
 QY 1149 TGTGCTTCTGCTGCTGCTGTAAGAGAGCTCTGAGATCTCGATGGCTTCAACAGTCC 1208
 DB 346 CySeGlyPheLeuLeuProMetLysGlnAlaProGlnLLeuSerGlySerAsnLysSer 365
 QY 1209 CTTCGCTTCATCTGATACAAACATCTGACATCCGCTTCAGTATTTGACCTTTGTG 1268
 DB 366 LeuAlaLeuHisLeuLLeuGlnAsnLLeuHisLLeuProGlnLTrnLysLeuVal 385
 QY 1269 GGTATGAAATATCTTCATGACAGACATCCCTGACCTGAAATCCGAGACAGTCTTGAGC 1328
 DB 386 AlaAsnGlnLysTrPheHisAspLysHisSerLeuThrGlnLLeuArgAspSerLeuLeuAsp 405
 QY 1329 TTGCTTGAGATGTGTCTTGTGTGCTCTGACATGATCAGAGTCAAGATCAAGAGAT 1388
 DB 406 LeuLeuGlyAspValPhePheValValProAlaLeuLLeuThrAlaArgTrpHisArg--- 424
 QY 1389 GCTGTGACACTGTCTACTTCTATGAGTTTCGACACCGGCTTCAGTCTTTGAAGACAGC 1448
 DB 424 ----- 424

QY 1449 AAGCCGGCTTTTGTGCAAAACCGACACGCTGATGAATGCCGTTGTTCGTGTGTC 1508
 DB 424 ----- 424
 QY 1509 TTCTCGAAGGGGACATTGTTATGTTTCAGAGAGCCACGAGAGAGAAATTACTGAC 1568
 DB 425 -----GluGlyAlaThrGluGlnGlyLysLeuLeuSer 435
 QY 1569 CGAAGATGATGAATPACTGGGCTACCTTGTCCGACCGGGAATCTTAATGGGAACGAC 1628
 DB 436 ArgLysMetMetLysTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 455
 QY 1629 CTGTCTGTGGCCGACTTATATCTGACTGAGAGTCACTCAGCTGATGGAACATG 1688
 DB 456 LeuSerLeuTrpProAlaTrnAsnLeuThrGlnGlnTrnLeuGlnLeuAspLeuAsnMet 475
 QY 1689 AGCTTCGACAGAGACTCAAGAAACCGGGGTGATTTTGAACACGACCATCCCTCTG 1748
 DB 476 SerLeuGlyAlaArgLeuLysGluProArgValAspPheTrpHisSerThrLLeuProLeu 495
 QY 1749 ATCTGTCTGCTCCGACATGCTCCAGAGTCTCTTCTTCTTAATTCTCTCTC 1808
 DB 496 LLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 515
 QY 1809 CTCAGCCTTCTTTTCTTTTGTGCTCCT 1838
 DB 516 LeuGlnProPhePhePheCysAlaPro 525
 RESULT 8
 US-10-114-270-196
 ; Sequence 196, Application US/10114270
 ; Publication No. US20040030110A1
 GENERAL INFORMATION:
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Malvankar, Uriel M.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Patnurejan, Meera
 ; APPLICANT: Liu, Zhaozhong
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Li, Li
 ; APPLICANT: Vernus, Corine
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Pena, Carol E.A.
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Burgess, Catherine E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Gangolli, Bsha A.
 ; APPLICANT: Taupier Jr., Raymond J.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Lieke, Mario W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Edinger, Shlomo R.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Rothenberg, Mark E.
 TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same
 FILE REFERENCE: 21402-322C
 CURRENT APPLICATION NUMBER: US/10/114,270
 PRIOR FILING DATE: 2002-11-27
 PRIOR APPLICATION NUMBER: 60/281,086
 PRIOR FILING DATE: 2001-04-03
 PRIOR APPLICATION NUMBER: 60/281,136
 PRIOR FILING DATE: 2001-04-03
 PRIOR APPLICATION NUMBER: 60/281,863
 PRIOR FILING DATE: 2001-04-05


```
/ PRIOR APPLICATION NUMBER: 60/281,906
/ PRIOR FILING DATE: 2001-04-05
/ PRIOR APPLICATION NUMBER: 60/282,020
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,930
/ PRIOR FILING DATE: 2001-04-10
/ PRIOR APPLICATION NUMBER: 60/282,934
/ PRIOR FILING DATE: 2001-04-10
/ PRIOR APPLICATION NUMBER: 60/283,512
/ PRIOR FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/283,710
/ PRIOR FILING DATE: 2001-04-13
/ PRIOR APPLICATION NUMBER: 60/284,234
/ PRIOR FILING DATE: 2001-04-17
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ SEQ ID NO 196
/ LENGTH: 581
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-114-270-196

Alignment Scores:
Pred. No.: 1,32e-225 Length: 581
Score: 2597.50 Matches: 497
Percent Similarity: 95.07% Conservative: 4
Best Local Similarity: 94.31% Mismatches: 17
Query Match: 67.08% Indels: 9
DB: 15 Gaps: 3

US-10-023-515-1 (1-2158) x US-10-114-270-196 (1-581)
QY 186 GGGGCTTCTCTGTAAGGCCACAGAGAACACAGGCTGGATGATTGAGGCAAGCA 245
Db 27 GlyProSerLalagluyProGlnArgSerThrArgLeuGlyTrpIleGlnGlySGln 46
QY 246 GTACTGTGCTGGGAAGCCCTGCGCTGGAAGTCTTCTCGAGATCCCTTGTGCT 305
Db 47 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheLala 66
QY 306 CCCCCCTGGAGATCCCTGCGATTACGAACCCGACAGCTGCATCGCCCTGGATAACTTG 365
Db 67 ProProLeuGlySerLeuArgPheThrAsnProGlnProLalSerProTrpAsnLeu 86
QY 366 CGAAGAACCACTCCCACTTAATTTGCTTCAGAACTCAGAGTGGCTGCTTAGAT 425
Db 87 ArgGlnAlaThrSerTyProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeuAsp 106
QY 426 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGACTGCTACCTG 485
Db 107 GlnHisMetLeuValValHisTyProLysPheGlyValSerGluAspCyLeuTyLeu 126
QY 486 AACATCTATGCGCCGCCACCGCCGATACAGGCTCCAAAGCTCCCGTGTGTGTTC 545
Db 127 AsnIleTyAlaProLalHisLalAspThrGlySerLysLeuProValLeuValTrpPhe 146
QY 546 CCAGAGAGTGCCTTCAAGACTGGCTCAGCTCCATCTTGAATGGGTCCGCCCTGGCTGCC 605
Db 147 ProGlyGlyAlaPheTyThrGlySerLalAserLlePheAspGlySerAlaLeuLala 166
QY 606 TATGAGGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
Db 167 TyGluAspValLeuValValValValGlnTyArgLeuGlyLlePheGlyPhePheThr 186
QY 666 ACATGGATGATCAGATGCTCCGAGGAGAACCGGCGCTTCAAGGACCGAGTGGCTGTGC 725
Db 187 ThrTrpAspGlnHisLalAspProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer 206
QY 726 TGGGTCCAGAAAGACATGAGTCTTCTCGTGGGAGACCCAGCTGTGACATCTTGGC 785
Db 207 TrpValGlnTyAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 226
QY 786 GAGTCCGCGGAGGACATAGTGTCTTAAGTCTTAATCTGCTCCATGCGCAAGGCTTA 845
```

```
Db 227 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 246
QY 846 TTCCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTACCTGGAGGCCCATGAT 905
Db 247 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyLeuGluAlaHisAsp 266
QY 906 TATGAGAAAGTGAAGACCTGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 965
Db 267 TyGluLysSerGluAspLeuGlnValAlaHisPheCySGlyAsnAsnAlaSerAsp 286
QY 966 TCTGAGGCCCTGCTGAGGTGCTGAGACAAACCCCTCAAGAGCTGCTGACCTCAGC 1025
Db 287 SerGlnAlaLeuLeuArgCyLeuArgThrLysProSerLysGluLeuThrLeuSer 306
QY 1026 CAGAAACAAAGTCTTTCATCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1085
Db 307 GlnTySerThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeu 326
QY 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1145
Db 327 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHis 346
QY 1146 GAGTGGGCTTCCGCTGGCCCTTGAAGAGAC-----GCTCCAGATCTCAGTGGCTCC 1199
Db 347 GluCySGlyPheLeuLeuProMetValArgIleLeuAlaValHisThrAlaThrProSer 366
QY 1200 AACAGTCCCTTGGCCCTC-----CATCTGAATCAAAACATCTGCACATC 1244
Db 367 AsnArgAspAlaAlaLeuAlaSerThrAlaGlyHisPheHisArgArg-----HisIle 384
QY 1245 CCGCTCATGATATTGGACCTTTGTGGCTAATGAATTAATCTTCATGACAGACCTCCCTGACT 1304
Db 385 ProProGlnTyLeuHisLysLeuValAlaAsnGlnTyPheHisAspLysHisSerLeuThr 404
QY 1305 GAAATCCGAGACAGTCTTGTGAGACTTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGT 1364
Db 405 GluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValAlaProAlaLeu 424
QY 1365 ATCAGAGCTGATATACACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1424
Db 425 IleThrLalaArgTyHisAspArgAlaGlyAlaProValTyPheTyGluPheArgHis 444
QY 1425 CCGCTTCAGAGCTTGAAGACAGAGCCGGCTTTGTGTGAAGCCGACCAAGCTGATGA 1484
Db 445 ArgProGlnCyPhePheGluAspThrLysProAlaPheValLysAlaAspHisLalAspGlu 464
QY 1485 GTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1544
Db 465 ValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGluGlyAla 484
QY 1545 ACGAGAGAGAGAGATTAATGAGCCGGAAGATGATGAATTAATCTGGGCTACCTTGTCTGA 1604
Db 485 ThrGlnGlnGlnLysLeuLeuSerArgLysMetCysTyTrpAlaThrPheLalaArg 504
QY 1605 ACCGGAAATCTTAATGGAACGACCTGTCTGTGGCAGCTTAAATCTGACGAGAG 1664
Db 505 ThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyAsnLeuThrGlnGln 524
QY 1665 TACCTCCAGCTGAGATTTGAACATGAGCTTGGACAGAGACTCAAGAACCGCGGGTGGAT 1724
Db 525 TyLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArgAsp 544
QY 1725 TTTTGAACCAACACATCCCC 1745
Db 545 ValTrpValThrGlyTyTrpPro 551

RESULT 9
US-10-381-898-2
; Sequence 2, Application US/10381898
; Publication No. US2004008687A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valda; BAUGHN, Mariah R.;
```


APPLICANT: BOROMSKY, Mark L.; DING, Li;
APPLICANT: DUGGAN, Brendan; ELIOT, Vicki S.;
APPLICANT: GANDHI, Ameena R.; GRIFPIN, Jennifer A.;
APPLICANT: HARPLIA, April J.A.; ISON, Craig H.;
APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
APPLICANT: LEE, Ernestine A.; LO, Dying Anna M.;
APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
APPLICANT: SANJANALA, Madhusudan M.;
APPLICANT: TANG, Y. Tom; TRIBOUDET, Catherine M.;
APPLICANT: CHAWLA, Nandinder K.; WALSH, Roderick T.;
APPLICANT: WARREN, Bridget; XU, Yuming;
APPLICANT: YANG, Junming; YAO, Monique; YUE, Henry
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-023 USN
CURRENT APPLICATION NUMBER: US/10/381, 898
CURRENT FILING DATE: 2003-10-17
PRIOR APPLICATION NUMBER: PCT/US01/30662
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/236, 947
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/238, 864
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/242, 323
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/247, 581
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249, 519
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252, 834
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/250, 567
PRIOR FILING DATE: 2000-11-30
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7473645CD1
US-10-381-898-2
Alignment Scores:
Pred. No.: 8,88e-225 Length: 618
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 66.85% Indels: 23
DB: 15 Gaps: 1
US-10-023-515-1 (1-2158) x US-10-381-898-2 (1-618)
QY 186 GGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGCA 245
DB 92 GTPProSeRAlaGluGlyPProGlnaGaaenThraGleuGlyTrrpLleInoGlyyGln 111
QY 246 GTCACTGTGCTGGGAAGCCCTGTGCTGGAACGTGTTCTCGAGTCCCTTGTGCTGCT 305
DB 112 ValThrValleuGlySerProValProValaenValPheleuGlyValProhehlaala 131
QY 306 CCCCCGCTGGGATCCCTGGGATTTACGAACCCGAGCCTGCAATCGCCCTGGGATTAATTG 365
DB 132 ProPProleuGlySerleuArpPheThraenProGlnProAlaSerProTrrpAspAenleu 151
QY 366 CGAGAAGCACTCCACCTAATTGCTGCTCGAAGTCAAGAGTGGGCTGCTTAAGT 425
DB 152 ArgGluAlaThrSerTrrpAspAenleuGlySerGlnAaenSerGluTrrpleuAenleuAsp 171
QY 426 CAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGACTGCTTAACCTG 485
DB 172 GlnHlAmeleuAlyValHisTrrpProlyBpHeGlyValSerGluAspCyBleuTrrleu 191

QY 486 AACATCTATGCGCTGCCACAGCCGATACAGGCTCAAGCTCCCGCTTGTGTTGCTTC 545
DB 192 AenlleTrrAlaProAlaHisAlaBpThrglySerlyBleuProValleuValTrrpHe 211
QY 546 CCAGAGGTGCTTCAAGACTGGCTCAAGCCTCATCTTTGATGGTCCGCTGCTGCC 605
DB 212 ProGlyGlyAlaPheleuThrglySerAlaSerllePheAspGlySerAlaAenlaala 231
QY 606 TATGAGAGCTGCTGTTGTGCTGCTCAAGTCCGCGCTAGAAATATTTGGTCTTCAAC 665
DB 232 TrrGluAspValleuValValValGlnTrraGleuGlylAepheGlyPhePheThr 251
QY 666 ACATGGGATCAAGACTGCTCCGGGAAGTGGGCTTCAAGACCAAGTGGCTGCTGCC 725
DB 252 ThrTrrpAspGlnHisAlaProGlyAenTrrAlaPheLyAspGlnValAlaAlaAenSer 271
QY 726 TGGGTCCAGAAAGACATGAGTTCTTGGGTGGGAACCCAGCTGTGACATCTTTGGC 785
DB 272 TrrValGlnLyAaenlleGluPhePheGlylAepProSerSerValThrrllePheGly 291
QY 786 GAGTCCGCGGAGCCATAGTGTCTTATGCTTATACGTCTCCATGGCCAAAGCTTA 845
DB 292 GluSerAlaGlyAlaAlaSerValSerSerleuIleleuSerProMetAlaLyGlyleu 311
QY 846 TTCACAAAGCCATGATGAGAGTGGGAGGCTCATCTCCCTTACCTGGAGGCCATGAT 905
DB 312 PheHlslyBAlaIleMeleGluSerGlyAlaAlaIlelleProTrrleuGlnAlaHisBp 331
QY 906 TATGAGAAAGTGAAGACCTCGAGGTGGTGGACATTTCTGTGTAAACATGCTCGAC 965
DB 332 TrrGlnLySerGluAspPheGlnValValAlaHisPheCyGlyAaenBnAlaSerAep 351
QY 966 TTTGAGGCTTGTGAGGTGCTGAGGACAAACCTCCCAAGAGCTGTGACCTCAGC 1025
DB 352 SerGlnAlaLeuAaGlyCyBleuArGThrrLyBProSerLyGlnleuAenThrrleuSer 371
QY 1026 CAGAAACAAAGCTTTTCACTGAGGTGGTGAATGGCTTTCTTCTTAATGAGCTTCA 1085
DB 372 GlnLyBThrrLySerPheThrraGlyAlaBpGlyAlaBpPheProAsnGluProleu 391
QY 1086 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATGATCGAGTCAATAACAC 1145
DB 392 AspLeuAenSerGlnLyAlaPheLyAlaIleProSerllelleGlyValaenBnHis 411
QY 1146 GAGTGTGCTTCTGCTGCTGCTTATGAAGAGGCTCCTGAGATCTCAAGTGGCTCCAAG 1205
DB 412 GluCyGlyPheleuAenProMet----- 419
QY 1206 TCCCTTGGCCCTCCATCTGATTAACAAACATCTCGACATCCGCGCTCAGATTTGGACCTT 1265
DB 420 -----HislleProGlnTrrleuHlsAen 428
QY 1266 GTGGCTAATGATATCTTCATGACAGACATCCCTGACTGAATCCAGACAGTCTTGTG 1325
DB 429 ValAlaAenGlnTrrPheHisBpLyBHisSerleuThrglnlleArgBpSerleuAen 448
QY 1326 GACTTGTGAGATGTGTTCTTTGTGTCCTGCACTGATCAAGCTGATATCAACA 1385
DB 449 AspLeuAenGlyAspValPhePheValAlaProAlaAenlleThrrAlaArgTrrHisArg 468
QY 1386 GATGTGTGAGCACTGCTCACTTCAATGAGTTTGGGCAACCGGCTCAAGCTTGTGAAGC 1445
DB 469 AspAlaGlyAlaProValTrrPheTrrGluBpPheArgHisArgBpGlnCyBpHeGlyAsp 488
QY 1446 ACGAAGCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTGTTCGTTGT 1505
DB 489 ThrLyBProAlaPheVallyAlaBpHisAlaBpGlnValAlaArgPheValPheGlyGly 508
QY 1506 GCTTTCCTGAAGGGGAGCACTTGTATGTTCAAGAGGACCAAGAGAGAGAAATTACTG 1565
DB 509 AlaPheleuLyGlyAspAlleValMeCPheGlnGlyAlaThrglnGlnGlylulBleuAen 528


```
FILE REFERENCE: SHIG FP02US006
CURRENT APPLICATION NUMBER: US/10/233, 933A
CURRENT FILING DATE: 2002-09-03
PRIOR APPLICATION NUMBER: JP2002-057908
PRIOR FILING DATE: 2002-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 542
TYPE: PRT
ORGANISM: Felis catus
US-10-233-933A-2

Alignment Scores:
Pred. No.: 3,62e-170 Length: 542
Score: 1985.00 Matches: 373
Percent Similarity: 83.21% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 51.27% Indels: 2
Gaps: 2

US-10-023-515-1 (1-2158) x US-10-233-933A-2 (1-542)

QY 186 GGGCCCTTCTGTCGAGGGCCACAGAGAAACAGAGCTGGATTCAGGGCAAGCA 245
DB 22 GlyProAlaIaAspAlaProValArgSerThrArgLeuGlyTyrValArgGlyIysLeuGln 41
QY 246 GTCACTGTCGTGGAGAGCCCTGCTGTGAAAGCTGTTCTCGAGAGTCCCTTCTGCT 305
DB 42 ThrThrValIleuGlySerThrValProValAsnMetPheLeuGlyIleProTyrIaIaA 61
QY 306 CCCCCTGGGGATCCCTGCGATTTACGAACCCGACGCTGCATGCGCCGGGATACCTG 365
DB 62 ProProLeuGlyProLeuArgPheLeuGlySerGlnProIysProAlaLeuProGlyAsnAspPhe 81
QY 366 CGAAGAGCCACCTCTTACCTTAATTTGTCCTCCAGAACTCAGAGTGCCTGCTTAGAT 425
DB 82 ArgAsnIaThrSerTyrProIysLeuCySPheGlnAspLeuGlyTyrPheValSerTyr 101
QY 426 CAACACATGCTCAAGTGCATTAACCCGAATTCGAGAGTGCAGAAAGCTGCTCTACCTG 485
DB 102 GlnIaIaValIleuIysValArgTyrProIysLeuGlyIaIaSerGlnAspCySPheLeuTyr 121
QY 486 AACATCTATGCGCCGACGCGGATACAGAGCTCCAGCTCCCGCTGCTGCTGCTGCTG 545
DB 122 AsnIleTyrAlaProAlaIaIaAspAsnGlySerAsnLeuProValMetValTyrPhe 141
QY 546 CCAGAGGTGCTTCAGAGACTGCTCAGCTTCATCTTGAATGGTCCGCTGCTGCTGCC 605
DB 142 ProGlyGlyAlaPheIysMetGlySerIaSerSerPheAspGlySerAlaIaIaIa 161
QY 606 TATAGAGACGTGCTGCTGCTGCTGCTGCTCAGTACGCGGCTAGAGAAATTTGGTTCTTCA 665
DB 162 TyrGlnAspValIleuIleValIThrThrGlnTyrArgLeuGlyIlePheGlyPheAsp 181
QY 666 ACATGGATACGATGCTCCGCGGGAACGCGGCTTCAGAGCCGAGGTGCTGCTGCTGCC 725
DB 182 ThrGlyAspGlyIuHAspArgGlyAsnTyrAlaLeuLeuAspGlnValAlaIaIaLeuThr 201
QY 726 TGGGTCAGAGAACATGAGTCTTCCGCTGGGGAACCCAGCTCTGTGACCATCTTTGGC 785
DB 202 TyrValArgAspAsnIleGlyPhePheGlyGlyAspProArgSerValThrIlePheGly 221
QY 786 GAGTCCCGGGAGCCATTAAGTGTCTTATAGTCTTATCTGCCAGGCCAAAGGCTTA 845
DB 222 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAsnGlyLeu 241
QY 846 TTCCACAAAGCCATCATGAGAGAGTGGGTGCATCATCCCTTACCTG---GAGGCCAT 902
DB 242 PheHisIysAlaIleMetGlnSerGlyValAlaIleLeuProLeuLeuMetArgProPro 261
QY 903 GATTATGAGAGAGTGAACCTGAGAGTGGTTCGACATTTCTGTGTTAACAATGCTCA 962
DB 962
```

```
DB 262 GlyAspGluArgGlyIysAspLeuGlnValIleuAlaArgIleCySPeGlyCySPeHisAlaSer 281
QY 963 GACTCTGAGGCCCTGCTGAGGCTGCTGAGCAAAACCTCCAGAGACTGCTGACCCCTC 1022
DB 282 AspSerAlaIaIaLeuLeuGlnCySPeLeuArgAlaIysProSerGlnGlnIleuMetAspIle 301
QY 1023 AGCCAGAAAACAAAGCTCTTTCATCTCGAGTGGTGAATGGTCTTCTTCTTAATGAGCCT 1082
DB 302 SerIysIysLeuThrPheSerIleProValIleAspAspPhePheProAspIuPro 321
QY 1083 CTAGATCTATTTGCTCAGAAAGCATTTAAAGCATTCCTTCCATCATGAGAGCTCAATAC 1142
DB 322 ValAlaLeuLeuThrGlnIysAlaPheAsnSerValProSerIleIleGlyValAsnAsn 341
QY 1143 CACGAGTGTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
DB 342 HisGlnCySPeAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyIysSerAsn 360
QY 1203 AAGTCCCTTGGCCCTCATCTGATACAAACATCTCCATCCAGCATCCGCTCAGATTTGGAC 1262
DB 361 ArgSerLeuAlaLeuTyrLeuValHisThrPheLeuAsnIleProThrGlnTyrIleuHis 380
QY 1263 CTGTGCTTAATGAATTAATCTTCATGACAGACATCTCCCTGACTGAATTCGAGACAGTCTT 1322
DB 381 LeuValAlaAspHisTyrPheTyrAsnIysHisSerProValGlnIleArgAspSerPhe 400
QY 1323 CTGACCTTGTGAGAGATGTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
DB 401 LeuAspLeuLeuGlyAspValLeuPheValProGlyValIaIaThrAlaArgTyrHis 420
QY 1383 AGAGATGCTGTGACCTGCTCTACTTCTATGAGTTTGGACCGGCTCAGTGCCTTTGAA 1442
DB 421 ArgAspAlaGlyAlaProValTyrPheTyrGlnPheGlnHisProProGlnIysLeuAsn 440
QY 1443 GACACGAAGCCGCTTTTGTCAAAAGCCGACCACTGATGAATCCGCTTGTGTGCT 1502
DB 441 AspThrArgProAlaPheValIysAlaAspHisSerAspGlnIleArgPheValPheGly 460
QY 1503 GTGCTCTTCTGAGAGGGGAGATTTGTTTCCAAAGGACGACGAGAGAGAGAAAGTTA 1562
DB 461 GlyAlaPheLeuIysGlyAspIleValMetPheGlnGlyAlaIaIaIaIaIaIaIaIaIa 480
QY 1563 CTGAGCCGGAAGATGATGAATTAATCTGAGCTACTTGTGCTCGAACCGGGAATCTATGGG 1622
DB 481 LeuSerAlaGlySPeMetMetArgTyrTyrAlaAsnPheAlaAspGlnIysAspProAsnIys 500
QY 1623 AACGACTGTCTGTGTGCGACGCTTATATCTGACTGACAGATACCTTCAGCTGACTTG 1682
DB 501 GlnGlyValProLeuTyrProAlaTyrThrGlnSerGlnGlnIysLeuIysLeuAspLeu 520
QY 1683 AACATGAGCTTCGACAGAGACTCAAGAGAACCGGCGGTGATTTTGGACAGCACCATC 1742
DB 521 SerValSerValGlyGlnIysLeuIysGlnGlnIaIaIaIaIaIaIaIaIaIaIaIaIa 540

RESULT 12
US-10-233-933A-4
Sequence 4, Application US/10233933A
Publication No. US20040214171A1
GENERAL INFORMATION:
APPLICANT: Yamashita, Teisuro
TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
FILE REFERENCE: SHIG FP02US006
CURRENT APPLICATION NUMBER: US/10/233, 933A
PRIOR APPLICATION NUMBER: JP2002-057908
PRIOR FILING DATE: 2002-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 542
TYPE: PRT
ORGANISM: Felis catus
```


NUMBER OF SEQ ID NOS: 110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 93
LENGTH: 356
TYPE: PRT
ORGANISM: Homo sapiens
US-10-451-168-93

Alignment Scores:

Pred. No.:	1,086-151	Length:	356
Score:	1780.00	Matches:	340
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	45.97%	Indels:	0
DB:	15	Gaps:	0

US-10-023-515-1 (1-2158) x US-10-451-168-93 (1-356)

819 ATACTGCTCCCATGCGCAAGGCTTATTCACAAAGCATCATGAGAGTGGGCGC 878
17 IleuSerProMetAlaIleuSerGlyLeuPheHisIleValIleMetGluSerGlyValAla 36
879 ATCATCCCTTACTGAGGCGCCATGATTAAGAAGAGTGAAGACCTGCAAGGTGGTGA 938
37 IleIleProTyLeuGluIleValIleHisAspTyrgIuIuSerGluAspLeuGlnValValAla 56
939 CATTTCTGTGTAAACAATGCGCTCAGACTTGAAGCCCTGCTGAGGCTCTGAGACAAA 998
57 HisPheCysGlyAsnAsnHisLeuSerAspSerGluAlaLeuLeuArgCysLeuAspTrpHis 76
999 CCGTCACAGAGGCTGCTGACCCCTGACCCGCAAGAAACAAGTCTTTCATCTGAGTGGTAT 1058
77 ProSerIleGlyIleuLeuLeuThrLeuSerGlnIuIuSerPheHisIleValValAla 96
1059 GGTGCTTCTTCTCTAATGAGCCTCTAATCTATTGCTCAGAAAGCATTTAAAGCAAT 1118
97 GlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnIuIuAlaPheValAlaIle 116
1119 CCTTCATCATCTGGAGTCAATACCAAGAGTGGCTTCTGCTGCTCTAATGAAGAGCT 1178
117 ProSerIleIleGlyValAsnAsnHisGlyCysGlyPheLeuLeuProMetCysIuAla 136
1179 CCGTAGATCTCTGAGGCTCCAAACAAGTCCCTTGCCTCATCTGATACAAACATCTCTG 1238
137 ProGluIleLeuSerGlySerAsnIuIuSerLeuAlaLeuHisIleuIleGlnIleu 156
1239 CACATCCCGCCTCAGTATTGTGACCTTGTGCTAAATGATATCTTCATGACAGCACTCC 1298
157 HisIleProProGlnTyIleuHisIleuValAlaAsnGluTyIlePheHisAspTrpHisSer 176
1299 CTGACTGAATCCGACAGCTTTCTGCACTTGTGCAATGTTGTTCTTTGTGCTCT 1358
177 LeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValValPro 196
1359 GCACGTGATCAGCGCTCGATATCAGAGAGTGTGGGACCTGTCTACTCTAATGAGTGT 1418
197 AlaLeuIleHisAlaArgTrpHisAspAspAlaGlyAlaProValAlaTrpPheTyrgIuPhe 216
1419 CGGACCGGCTCAGTGTCTTTGAGACAGCAAGCCGGCTTTTGTGCAAGCCGACCAAGCT 1478
217 ArgHisAspProGlnCysPheGluAspTrpIuIuSerProAlaPheValIleAspHisVala 236
1479 GATGAAGTCCGCTTGTGTGTGGTGTGCTTCTGAGAGGGGACATTTATTTGTCGA 1538
237 AspGluValArgPheValPheGlyGlyAlaPheLeuIuIuSerGlyAspIleValMetPheGlu 256
1539 GGAGCGACGAGAGAGAGAAAGTACAGAGCCGAGAGATGATGAATATCTGGGCTACCTT 1598
257 GlyAlaTrpTrpGlnGlnIuIuSerLeuSerAspTrpGlySerMetCysTrpTrpAlaTrpHis 276
1599 GCTCGAACCGGAGATCTAATGGAAGCACTGTCTCTGTGCGCAGCTTATATCTGACT 1658
277 AlaArgTrpTrpGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeuThr 296

1659 GAGCAGTACCTTCAGCTGATGTAACATGAGCCTCGGACAGAGACTCAAAAGACCGCG 1718
297 GluGlnTyIleLeuGlnIleuAspLeuAsnMetSerLeuGlyGlnArgIleuIuSerProArg 316
1719 GTGATTTTGGACCAAGCAACATCCCTGATCTGTCTGTGCTCGACATGCTCCACT 1778
317 ValAspPheTrpTrpHisSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHisSer 336
1779 CCTCTTCTCTTAACTTTCTCTCTCTCTCCACGCTTTTCTTTTCTTTGTGCTCT 1838
337 ProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePhePheCysAlaPro 356
RESULT 14
US-09-925-298-689
Sequence 689, Application US/09925298
Publication No. US20020039764A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: P0103
CURRENT APPLICATION NUMBER: US/09/925,298
PRIORITY FILING DATE: 2001-08-10
PRIORITY APPLICATION NUMBER: PCT/US00/05881
PRIORITY FILING DATE: 2000-03-08
PRIORITY APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 689
LENGTH: 549
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (7)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-689
Alignment Scores:
Pred. No.: 8,186-102 Length: 549
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 31.77% Indels: 24
Gaps: 8
US-10-023-515-1 (1-2158) x US-09-925-298-689 (1-549)
186 GGGCTTGTGCTGAAGGCGCACAGAGAACACAGGCTGGATGATTCAGGCAAGCA 245
16 GlyGlnAspSerAlaSerProIleArgTrpHisTrpHisGlnValLeuIuSerLeu 35
246 GTCACTGTGCTGGAGAGCCCTGTGCTGAACGTTCTCTGGAGTCCCTTGTGCT 305
36 ValHisValIuIuSerGlyAlaAsnAlaGlyAlaGlnTrpPheLeuGlyIleProPheAlaIu 55
306 CCGCGCTGGAGTCCCTGATTTTACGAACCGGACGCTGATCGCCCTGGGATTAATG 365
56 ProProLeuGlyIuIuSerGlnAlaProProGluIuProGluIuSerTrpSerGlyVal 75
366 CGAGAAGCACTCTTACCTTAATTTGTGCTCCAGAAC-----TCAGAG 410
76 ArgAspGlyTrpTrpHisProAlaMetCysLeuGlnIleuTrpAlaValGluSerGlu 95
411 TGGCTGCTTAAATGATCAACATGCTCAAGGTCATTACCGGAATTCGAGAGTCAGAA 470
96 PheLeu-----SerGlnPheAsnMetTrpPheProSerAspSerMetSerGlu 111
471 GACTGCTTACCTGAACATGATGAGCTGCGCTGCCAGCGGATGACAGGCTCAAGCTCC 530

```

Db      112 AspCysLeuTyrLeuSerIleTyrThrProAlaHisSerHisGlnGlySerAsnLeuPro 131
QY      531 GTCTTGATGTTGTTCCACAGAGTGCCTTCAAGACTGGCTGACCTTCATCTTGATGGG 590
Db      132 ValMetValTyrIleHisGlyGlyAlaLeuValPheGlyMetAlaSerLeuTyrAspGly 151
QY      591 TCCGCGCTGGCTGATGAGGACGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 650
Db      152 SerMetLeuAlaAlaLeuGlnHisValValValIleIleGlnTyrArgLeuGlyVal 171
QY      651 TTGTGTTCTTCCACATGGAGTCAAGCATGAGTCTCCGGGAACTGGGCTTCAAGACAG 710
Db      172 LeuGlyPhePheSerThrGlyAspArgHisAlaThrGlyAsnTyrGlyLeuAspGln 191
QY      711 GTGGTGTCTGTCTGCTGGTCCAGAAAGACATCGAGTTCTTGGTGGGAGCCCAAGCTCT 770
Db      192 ValAlaAlaLeuAlaGlyTyrValGlnGlnHisIleAlaHisPheGlyGlyAsnProAspArg 211
QY      771 GTGACCATCTTGGCGAGTCCGCGGAGCCATAAGTGTCTTACTTATATACGTCTCC 830
Db      212 ValThrIlePheGlyGlnSerAlaGlyGlyThrSerValSerSerLeuValValSerPro 231
QY      831 ATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGGCTGACATCCCTTAC 890
Db      232 IleSerGlnGlyLeuPheHisGlyAlaIleMetGlySerGlyValAlaLeuLeuProGly 251
QY      891 CTGAGAGCCCATGATGTAGAAAGAGTGAAGACCTG-----CAGGTGTTGACATTC 944
Db      252 LeuIleAla-----SerSerAlaAspValIleSerThrValAlaAlaAsnLeu 267
QY      945 TGTGTGACAAATGCGTCACTGAGTGAAGCCCTGCTGAGGTGCTGAGACAAACCTCC 1004
Db      268 SerAlaCysAspGlnValAspSerGlnAlaLeuValGlyCysLeuArgGlySerIlys 287
QY      1005 AAGAGCTGCTGACCTCAGCCAGCAAGAAACAAAGTCTTTCATCGAGTGGTGTGCT 1064
Db      288 GlnGlnIleLeuAlaIleAsnIysProPheIysMetIleProGlyValAlaAspGlyVal 307
QY      1065 TTCTTCTTAATGAGCTCTAGACTTATGTCTCAGAAAGCATTTAAGCAATTCCTCC 1124
Db      308 PheLeuProArgHisProGlnGlnLeuAlaSerAlaAspPheGlnProValProSer 327
QY      1125 ATCATCGAGTCAATTAACAGAGTGGCTTCCGCTCCCT-----ATGAAG 1172
Db      328 IleValGlyValAsnAsnHisGlnPheGlyTyrPheIleProIysValMetArgIleTyr 347
QY      1173 GAGGCTCTGAGATCTCAGTGGCTCCAAACAAGTCCCTTGCCTCATCTGATACAAAC 1232
Db      348 AspThrGlnIlyGlnMetAspArgGlnIleSerGlnAlaAlaLeuGlnIlyMetLeuThr 367
QY      1233 ATCTGACATACCCGCTCAGATTAATGGACCTTGGGCTAATGAATCTTCATGACAAAG 1292
Db      368 LeuLeuMetLeuProProThrPheGlyAspLeuAlaGlnGlyTyrIleGlyAspAsn 387
QY      1293 CACTCCCTGATGAATCCGAGACAGTCTTCTGACCTTGGAGATGTTCTTTGTG 1352
Db      388 GlyAspProGlnThrLeuGlnAlaGlnPheGlnIleMetMetAlaAspSerMetPheAla 407
QY      1353 GTCCCTGACATGATACAGCTCGATATACAGAGATGCTGTGACCTGTCTACTTCTAT 1412
Db      408 IleProAlaLeuGlnValAlaHisPhe---GlnCysSerArgAlaProValTyrPheTyr 426
QY      1413 GAGTTTCGGACACCGGCTCAGTGTCTTGAAGACAGACCGGCTTTTGTCAAGCCGAC 1472
Db      427 GlnPheGlnHisGlnProSerTyrPheIlyAsnIleArgProProHisMetIlyAlaAsp 446
QY      1473 CACGCTATGAAGTCCGCTTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1532
Db      447 HisGlyAspGlnLeuProPheValPhe---ArgSerPhePheGlyGlyAsnTyrIleIys 465
QY      1533 TTCGAAGAGACGACGAGAGAGAGAGTACTGAGCCGGAAGATGAGAAATCTGGGCT 1592

```

```

Db      466 Phe-----ThrGlnGlnGlnGlnLeuSerArgIysMetMetIysTyrTrpAla 482
QY      1593 ACCTTGCTTCGAACCGGGAATCCTAATGGGAACGACCTGTCTGTGGCCAGCTTAAAT 1652
Db      483 AsnPheAlaArgAsnIlyAsnProAsnGlnGlyGlyLeuProHisTyrProLeuPheAsp 502
QY      1653 CTGACTGACAGTACTCCTCAGCTGAGCTTGAACATGACGCTTGGACAGACTCAAGAA 1712
Db      503 GlnGlnGlnGlnIlyLeuGlnLeuAsnLeuGlnProAlaValGlyArgAlaLeuIysAla 522
QY      1713 CCGCGGTGATTTTGGACGACGACCATCCCC 1745
Db      523 HisArgLeuGlnPheTyrIlyValAlaLeuPro 533

RESULT 15
US-10-102-806-689
/ Sequence 689, Application US/10102806
/ Publication No. US20030054421A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: P4103P1C1
/ CURRENT APPLICATION NUMBER: US/10/102,806
/ PRIOR APPLICATION NUMBER: 09/925,298
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05881
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 846
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 689
/ LENGTH: 549
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (7)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-689

Alignment Scores:
Pred. No.: 8,18e-102 Length: 549
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatch: 168
Query Match: 31.77% Indels: 24
DB: 14 Gaps: 8

US-10-023-515-1 (1-2158) x US-10-102-806-689 (1-549)
QY      186 GGGCCCTTCTGTAAGAGGCGCACAGAGAAACACAGCTGGAGTGAATTCAGGCAAGCA 245
Db      16 GlnGlnAspSerAlaSerProIleArgThrHisThrGlnIlyGlnValLeuGlySerLeu 35
QY      246 GTCACTGTGCTGGGAAGCCCTGTGCTGTAACGTGTCTCGAGAGTCCCTTGTGCT 305
Db      36 ValHisValIlyGlyAlaAsnAlaGlyAlaGlnThrPheLeuGlyIleProPheAlaIys 55
QY      306 CCCCCGCTGGATCCCTGCAATTAACGAACCGGAGCCTGATGCCCTGGGATTAATCG 365
Db      56 ProProLeuGlyProLeuArgPheAlaProGlnProProGlnSerTyrPserGlyVal 75
QY      366 CGAAGACCACTCTTACCTTAATTTGTGCTCGAAGC-----TCAGAG 410
Db      76 ArgAspGlyThrThrHisProAlaMetCysLeuGlnAspLeuThrAlaValGlnSerGln 95
QY      411 TGGCTGCTTAAATCAACACATGCTCAAGTGAATTAACCGGAATTCGAGTCAAGAA 470

```

```

Db      PheLeu-----SerGlnPheAenMetThrPheProSerAspSerMetSerGln 111
QY      471 GAGTGGCTTACCTGAATCATATGAGCGCTGGCCACGGCGATACAGGCTCCAAAGCTCCCC 530
Db      112 AaPcYbLeuYrLeuSerIleYrThrProAlaHisSerHisGlnGlySerAenLeuPro 131
QY      531 GTCTTGATGTGGTCCCAAGAGTGCCTTCAAGACTGGCTCAGCTCCATCTTTGATGGG 590
Db      132 ValMetValTTrpIleHisGlyValAlaLeuValPheGlyMetAlaSerLeuYrTrpGly 151
QY      591 TCCGCGCTGGCTGCTATGAGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 650
Db      152 SerMetLeuAlaAlaLeuGlnLysAenValValValIleIleGlnIlyrArgLeuGlyVal 171
QY      651 TTTGGTTTCTTCAACCATGGATCAGCATGTCTCCGGGGAAGTGGGCGCTTCAAGACCG 710
Db      172 LeuGlyPhePheSerThrGlyAspLysHisAlaThrGlyAsnTrpGlyYrLeuAspGln 191
QY      711 GTGGCTGCTGTGCTGGGTCCAGAAACATCGATGTTTCCGTGGGAGCCCGACGCTCT 770
Db      192 ValAlaAlaLeuAArgTrpValGlnGlnAenIleAlaHisPheGlyGlyAsnProAspArg 211
QY      771 GTGACCATCTTTGGGAGTCCGGGAGGCGCATAGTGTCTTACTGTCTCTCC 830
Db      212 ValThrIlePheGlyLeuSerAlaGlyThrSerValSerSerLeuValValSerPro 231
QY      831 ATGGCCAAAGGCTTATTCACAAACCATCAGAGAGTGGGTGGCCATCATCCCTTAC 890
Db      232 IleSerGlnGlyLeuPheHisGlyValIleMetGlnSerGlyValAlaLeuLeuProGly 251
QY      891 CTGGAGGCGCCATGATTATGAGAAAGTGAGACGTG-----CAGGTGGTGACATTTTC 944
Db      252 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAenLeu 267
QY      945 TGTGGTAAACAATGCGCTCAGACTCTGAGGCGCTGAGAGTGCCTGAGGACAAACCTTCC 1004
Db      268 SerAlaCyAspGlnValAspSerGlnAlaLeuValGlyCysLeuAArgGlyLysSerLys 287
QY      1005 AAGAGAGTGTCTGACCTCAGCCACGACAAACAAAGTCTTTCATCTCAGAGTGGTGTGCT 1064
Db      288 GlnGlnIleLeuAlaIleAsnLysProPheLysMetIleProGlyValValAspGlyVal 307
QY      1065 TTCTTTCCATATGAGCTCTATGATCTATGTCTCAGAAAGCATTTAAAGCATTCCTTCC 1124
Db      308 PheLeuProAlaGlnIleProGlnIleuLeuAlaSerAlaAspPheGlnProValIleProSer 327
QY      1125 ATCATCGAGATGAATTAACCAAGAGTGTGGCTTCTGTGCTCT-----ATGAG 1172
Db      328 IleValGlyValAlaAsnAenGlnPheGlyTrpLeuIleProLysValMetAArgIleYr 347
QY      1173 GAGGCTCCTGAGATCTCTAGTGGCTTCCAACAAAGTCCCTTGCTCATCTGATACAAAC 1232
Db      348 AspThrGlnLysGlnMetAspAArgLysAlaSerGlnAlaAlaLeuGlnLysMetLeuThr 367
QY      1233 ATCTCGACATCCCGCTCAGATTTTGACACTGTGGGTAAATGAATCTTCATGACAAAG 1292
Db      368 LeuLeuMetLeuProThrPheGlyAspLeuLeuAArgGlnLysTrpIleGlyAspAen 387
QY      1293 CACTCCCTGACTGAATCCAGACAGATCTTCTGACTTGTGAGATGTGTCTTCTTGTG 1352
Db      388 GlyAspProGlnThrLeuGlnAlaGlnPheGlnLysMetMetAlaAspSerMetPheVal 407
QY      1353 GTCCCTGACATGATACAGCTCGATATACAGAGATGCTGTGACACTGTCTTACTTAT 1412
Db      408 IleProAlaLeuGlnValAlaHisPhe--GlnCysSerAArgAlaProValYrPheYr 426
QY      1413 GAGTTTGGGCAACGGGCTCAGAGCTTTGAAACACAGACCGGCTTTGTCAAGCGGAC 1472
Db      427 GlnPheGlnHisGlnProSerTrpLeuLysAsnIleArgProProHisMetLysAlaAsp 446
QY      1473 CAGCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1532
Db      447 HisGlyAspGlnLeuProPheValPhe--ArgSerPhePheGlyGlyAsnYrIleLys 465

```

```

QY      1533 TTGAAAGAGCCACGAGAGAGGAAGTTTACTGACCCGGAAGATGATGAATATATGAGCT 1592
Db      466 Phe-----ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 482
QY      1593 ACCTTGTCTGAACCGGAAATCTTAATGGAAACGACTGTCTGTGGCCAGCTTATAT 1652
Db      483 AsnPheAlaArgAenGlyAsnProAenGlyGlnGlyLeuProHisIleTrpProLeuPheAsp 502
QY      1653 CTGACTGAGCACTACCTCAGCTGAGCTTGAACATGAGCGCTGGACAGACTCAAGAA 1712
Db      503 GlnGlnGlnGlnLysLeuGlnLeuAenGlnProAlaValGlyAArgAlaLeuLysAla 522
QY      1713 CCGGAGTGAATTTTGGACCCAGCACCATCC 1745
Db      523 HisArgLeuGlnPheTrpLysAlaLeuPro 533

```

Search completed: June 15, 2005, 09:36:39
Job time : 181.194 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 22:52:56 ; Search time 1123.45 Seconds
(without alignment)
9634.209 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 1746

Sequence: 1 atgcacagcgagctactctc.....ttctcttgcctctga 1746

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 309530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
2	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
3	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
4	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
5	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
6	1652.8	94.7	1728	US-10-451-168-40	Sequence 45, Appli
7	1651.6	94.6	2232	US-10-433-256-23	Sequence 23, Appli

8	1631.2	93.4	2092	17	US-10-104-047-249	Sequence 249, App
9	1435.6	82.2	1746	18	US-10-114-270-195	Sequence 195, App
10	1432.4	82.0	1746	18	US-10-451-168-41	Sequence 41, Appli
11	1386.2	75.4	1857	18	US-10-381-898-20	Sequence 20, Appli
12	1344.4	77.0	1962	17	US-10-094-749-736	Sequence 736, App
13	1099	62.9	2145	20	US-10-233-933A-3	Sequence 3, Appli
14	1063.2	60.9	1629	20	US-10-233-933A-1	Sequence 1, Appli
15	1021.4	58.5	1071	18	US-10-451-168-42	Sequence 42, App
16	508.6	29.1	965	18	US-10-302-172-155	Sequence 155, App
17	406.6	23.3	2191	9	US-09-880-107-3854	Sequence 3854, Ap
18	406.6	23.3	2191	9	US-09-954-531-1038	Sequence 1038, Ap
19	406.6	23.3	2191	10	US-09-873-367C-828	Sequence 828, App
20	406.6	23.3	2191	10	US-09-873-367C-829	Sequence 829, App
21	406.6	23.3	2191	20	US-10-858-271-27	Sequence 27, Appli
22	406.6	23.3	2191	21	US-10-843-641A-828	Sequence 828, App
23	406.6	23.3	2191	21	US-10-843-641A-829	Sequence 829, App
24	406.6	23.3	2191	21	US-10-843-641A-2105	Sequence 2105, Ap
25	406.6	23.3	2484	14	US-10-102-806-271	Sequence 271, App
26	406.6	23.3	2484	14	US-10-102-806-271	Sequence 271, App
27	400.2	22.9	2092	21	US-10-956-157-4153	Sequence 4153, Ap
28	400.2	22.9	3824	13	US-09-931-836-22	Sequence 22, Appli
29	400.2	22.9	3824	13	US-10-036-342-22	Sequence 22, Appli
30	400.2	22.9	3824	13	US-10-036-041-22	Sequence 22, Appli
31	400.2	22.9	3824	14	US-10-028-072-541	Sequence 541, App
32	400.2	22.9	3824	14	US-10-035-855-22	Sequence 22, Appli
33	400.2	22.9	3824	14	US-10-140-808-541	Sequence 541, App
34	400.2	22.9	3824	14	US-10-121-049-541	Sequence 541, App
35	400.2	22.9	3824	14	US-10-123-904-541	Sequence 541, App
36	400.2	22.9	3824	14	US-10-140-470-541	Sequence 541, App
37	400.2	22.9	3824	14	US-10-175-746-541	Sequence 541, App
38	400.2	22.9	3824	14	US-10-176-918-541	Sequence 541, App
39	400.2	22.9	3824	14	US-10-176-918-541	Sequence 541, App
40	400.2	22.9	3824	14	US-10-227-864-209	Sequence 209, App
41	400.2	22.9	3824	14	US-10-036-214-22	Sequence 22, Appli
42	400.2	22.9	3824	14	US-10-137-865-541	Sequence 541, App
43	400.2	22.9	3824	14	US-10-140-474-541	Sequence 541, App
44	400.2	22.9	3824	14	US-10-035-719-22	Sequence 22, Appli
45	400.2	22.9	3824	14	US-10-142-431-541	Sequence 541, App

ALIGNMENTS

RESULT 1
US-10-023-515-3
; Sequence 3, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279, 508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-515-3

Query Match 100.0%; Score 1746; DB 13; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGCACAGGAGCTACTGCTCATCTGCTTCACATGATGCTTTCTCTGATTCACAGCCC 60
|||||

Db 1 ATGCCACAGGGAGCTTACTTCATCTGCTTCACAAAGTGCTTTTCTCGATTCCTCAGCCC 60
QY 61 CTGTTGGGACACAGACAGTGGGGAAAACTGGCCCTTCTGCTGAAGGCCCAAGAGAAC 120
Db 61 CTGTTGGGACACAGACAGTGGGGAAAACTGGCCCTTCTGCTGAAGGCCCAAGAGAAC 120
QY 121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTG 180
Db 121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTG 180
QY 181 AACGGTCTCTGGAGTCCCTTTGCTGCTCTCCCGCTGGGATCCCTGCAATTTACGAAC 240
Db 181 AACGGTCTCTGGAGTCCCTTTGCTGCTCTCCCGCTGGGATCCCTGCAATTTACGAAC 240
QY 241 CCGAGAGCTGATGACCTGGGATTAATTGGAGAGGCACTCTCTAACCCTAATTTGTGC 300
Db 241 CCGAGAGCTGATGACCTGGGATTAATTGGAGAGGCACTCTCTAACCCTAATTTGTGC 300
QY 301 CTCGAGAACTCAGAGTGGCTGCTTATGATCAACAATGCTCAAGGTGCAATTAACCGAAA 360
Db 301 CTCGAGAACTCAGAGTGGCTGCTTATGATCAACAATGCTCAAGGTGCAATTAACCGAAA 360
QY 361 TTGCGAGTGTCAAGAACTGCTCTACCTGAACAATGAGCGCTGCGCCACGCCGATACA 420
Db 361 TTGCGAGTGTCAAGAACTGCTCTACCTGAACAATGAGCGCTGCGCCACGCCGATACA 420
QY 421 GGGCTCCAGCTCCCGCTTGTGTGTGTGTTCCAGAGAGTGCCTTCAAGAATGCTGAGCC 480
Db 421 GGGCTCCAGCTCCCGCTTGTGTGTGTGTTCCAGAGAGTGCCTTCAAGAATGCTGAGCC 480
QY 481 TCCATCTTTATGATGGATCCGCTGGCTGCTATGAGAGAGTGTGTGTGTGTGTGTGCAAG 540
Db 481 TCCATCTTTATGATGGATCCGCTGGCTGCTATGAGAGAGTGTGTGTGTGTGTGTGCAAG 540
QY 541 TACCGGCTAGGAATATTGGTTTCTTACCAACATGAGGATCAAGATCTCCGGGAACTGG 600
Db 541 TACCGGCTAGGAATATTGGTTTCTTACCAACATGAGGATCAAGATCTCCGGGAACTGG 600
QY 601 GCGTTAAAGAACCAAGTGTGCTGTCTGTGCTGGGCTCAGAGAAACATGCAATTTCTGAGT 660
Db 601 GCGTTAAAGAACCAAGTGTGCTGTCTGTGCTGGGCTCAGAGAAACATGCAATTTCTGAGT 660
QY 661 GGGGAGCCCAAGCTGTGTGACCACTTTTGGGAGTCCGGGGGAGCCATAAGTCTTCTAGT 720
Db 661 GGGGAGCCCAAGCTGTGTGACCACTTTTGGGAGTCCGGGGGAGCCATAAGTCTTCTAGT 720
QY 721 CTTATATCTGTCTCCATGAGCCAAAGCTTATTCACAAAGCCATGAGAGTGGGGTG 780
Db 721 CTTATATCTGTCTCCATGAGCCAAAGCTTATTCACAAAGCCATGAGAGTGGGGTG 780
QY 781 GGCATCATCCCTTACCTGAGGCGCATGATTAATGAGAGAGTGTGAGGATCTGAGTGGTT 840
Db 781 GGCATCATCCCTTACCTGAGGCGCATGATTAATGAGAGAGTGTGAGGATCTGAGTGGTT 840
QY 841 GCACATTTCTGTGTAAACATGCGTCAAGCTGTGAGGCTGCTGAGGTGCTTAAGACA 900
Db 841 GCACATTTCTGTGTAAACATGCGTCAAGCTGTGAGGCTGCTGAGGTGCTTAAGACA 900
QY 901 AAACCCCTCAAGAGCTGCTGACCTTCAAGCCAGAAAACAAAGTCTTTCACGAGTGGTT 960
Db 901 AAACCCCTCAAGAGCTGCTGACCTTCAAGCCAGAAAACAAAGTCTTTCACGAGTGGTT 960
QY 961 GATGGTGTCTTCTTCTTAATGAGCTTATGATCTATGCTGAGAAAGCATTTAAAGCA 1020
Db 961 GATGGTGTCTTCTTCTTAATGAGCTTATGATCTATGCTGAGAAAGCATTTAAAGCA 1020
QY 1021 ATTCTCTTCATCATGAGATCAATTAACAGAGTGTGGCTTCTGCTGCTATGAAGAG 1080
Db 1021 ATTCTCTTCATCATGAGATCAATTAACAGAGTGTGGCTTCTGCTGCTATGAAGAG 1080
QY 1081 GCTCTGAGATCTCAGTGGCTCCAAAGAGTCCCTTGGCTTCATCTGATACAAACATCT 1140
Db 1081 GCTCTGAGATCTCAGTGGCTCCAAAGAGTCCCTTGGCTTCATCTGATACAAACATCT 1140

QY 1141 CTGACATCCCGGCTGAGTATTGGACCTTGTGGCTTAATGAATATTCCATGACAGAC 1200
Db 1141 CTGACATCCCGGCTGAGTATTGGACCTTGTGGCTTAATGAATATTCCATGACAGAC 1200
QY 1201 TCCCTGATGAAATCCGAGACAGTCTTCTGAGCTTGTGAGATGTGTTCTTGTGCTC 1260
Db 1201 TCCCTGATGAAATCCGAGACAGTCTTCTGAGCTTGTGAGATGTGTTCTTGTGCTC 1260
QY 1261 CCGACATGATCAAGCTGATATCAAGAGTGTGTGCTGCTCTGATCTTCTATGAG 1320
Db 1261 CCGACATGATCAAGCTGATATCAAGAGTGTGTGCTGCTCTGATCTTCTATGAG 1320
QY 1321 TTTGGACACCGGCTCAGTCTTTGAGACAGAGCCGGCTTTTGTCAAGCCGACAC 1380
Db 1321 TTTGGACACCGGCTCAGTCTTTGAGACAGAGCCGGCTTTTGTCAAGCCGACAC 1380
QY 1381 GCTGATGAATCCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 1381 GCTGATGAATCCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
QY 1441 GAGGAGCCAGAGAGAGAGAGATTAAGTGAAGCCGAGAGATGAATTAAGTGAAG 1500
Db 1441 GAGGAGCCAGAGAGAGAGAGATTAAGTGAAGCCGAGAGATGAATTAAGTGAAG 1500
QY 1501 TTTGCTCAACCGGGAATCTTAATGGGACGACCTGTCTGTGTGCGCAGCTTATATCTG 1560
Db 1501 TTTGCTCAACCGGGAATCTTAATGGGACGACCTGTCTGTGTGCGCAGCTTATATCTG 1560
QY 1561 ACTGAGCATCTCCAGCTGAGCTTGAACATGAGCCTCGGACAGAGATCTCAAGAAC 1620
Db 1561 ACTGAGCATCTCCAGCTGAGCTTGAACATGAGCCTCGGACAGAGATCTCAAGAAC 1620
QY 1621 CGGGTGAATTTTGGACCAAGCAATCCCTGATCTGTGCTGCTCCAGATGCTCCAC 1680
Db 1621 CGGGTGAATTTTGGACCAAGCAATCCCTGATCTGTGCTGCTCCAGATGCTCCAC 1680
QY 1681 AGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
Db 1681 AGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1740
QY 1741 CTTTGA 1746
Db 1741 CTTTGA 1746

RESULT 2
US-10-674-636-3
Sequence 3, Application US/10674636
Publication No. US20040086922A1
GENERAL INFORMATION:
APPLICANT: Curtiss, Roy A. J.
APPLICANT: Siles-Santiago, Imaculada
TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
FILE REFERENCE: 10448-122001
CURRENT APPLICATION NUMBER: US/10/674,636
PRIOR FILING DATE: 2003-09-29
PRIOR APPLICATION NUMBER: US/10/023,515
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256,369
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/279,508
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1746
TYPE: DNA
ORGANISM: Homo sapiens
US-10-674-636-3

Query Match 100.0%; Score 1746; DB 18; Length 1746;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	ATGCCAAGAGGACTTACTTCAATCTCTCTTCACAATAGTGTCTTTTCTGATTTCTCAGGCC	60
OY	1	ATGCCAAGAGGACTTACTTCAATCTCTCTTCACAATAGTGTCTTTTCTGATTTCTCAGGCC	60
Db	1	ATGCCAAGAGGACTTACTTCAATCTCTCTTCACAATAGTGTCTTTTCTGATTTCTCAGGCC	60
OY	61	CTGTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGTGAAGGGCCACAGAGAAC	120
Db	61	CTGTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGTGAAGGGCCACAGAGAAC	120
OY	121	ACCAGGCTGGGATGGAATTCAGGGCAAGCAATCTGTGCTGGGAAAGCCCTGTGCTGTG	180
Db	121	ACCAGGCTGGGATGGAATTCAGGGCAAGCAATCTGTGCTGGGAAAGCCCTGTGCTGTG	180
OY	181	AACGTGTCTCTCGAGAGTCCCTTTGCTGTGCTCCCGCTGGGATGCCCTGCAATTTACGAAC	240
Db	181	AACGTGTCTCTCGAGAGTCCCTTTGCTGTGCTCCCGCTGGGATGCCCTGCAATTTACGAAC	240
OY	241	CCGACGCTGCATGCGCCCTGGGATTAATTTGCGAGAAAGCACTCTCTTAATTTGTGTC	300
Db	241	CCGACGCTGCATGCGCCCTGGGATTAATTTGCGAGAAAGCACTCTCTTAATTTGTGTC	300
OY	301	CTCCGAATCTCAGAGTGTGCTCTTAATATCAACAATGCTCAAGGTGATTAATCCGAAA	360
Db	301	CTCCGAATCTCAGAGTGTGCTCTTAATATCAACAATGCTCAAGGTGATTAATCCGAAA	360
OY	361	TTCCGAGTGTCAAGAACTGCTCTACCTGAACATCTAATGGCGCTGCCACGCGCATACA	420
Db	361	TTCCGAGTGTCAAGAACTGCTCTACCTGAACATCTAATGGCGCTGCCACGCGCATACA	420
OY	421	GAGCTTCAAGCTTCCCGCTTGTGTGTGTCCAGAGAGGTGCCCTTCAAGCTGGCTCAGCC	480
Db	421	GAGCTTCAAGCTTCCCGCTTGTGTGTGTCCAGAGAGGTGCCCTTCAAGCTGGCTCAGCC	480
OY	481	TCCATCTTTGATGGGTCCGGCTCTGGCTGTGATGAGACATGTGTGTGTGTGTGTGTGT	540
Db	481	TCCATCTTTGATGGGTCCGGCTCTGGCTGTGATGAGACATGTGTGTGTGTGTGTGTGT	540
OY	541	TACCGGCTAGGAAATATTTGATTTCTTCAACACAATGAGATTCAGCATGCTCCGGGAAACTGG	600
Db	541	TACCGGCTAGGAAATATTTGATTTCTTCAACACAATGAGATTCAGCATGCTCCGGGAAACTGG	600
OY	601	GCTTTCAGAGACCAAGTGGCTGCTGTCTCTGGGTCCAGAAACATCGAATTTCTTCGGT	660
Db	601	GCTTTCAGAGACCAAGTGGCTGCTGTCTCTGGGTCCAGAAACATCGAATTTCTTCGGT	660
OY	661	GGGAGCCCCAGCTCTGTGACCACTTTTGCGAGGTCCGGGAGCCATTAAGTGTCTTGTGT	720
Db	661	GGGAGCCCCAGCTCTGTGACCACTTTTGCGAGGTCCGGGAGCCATTAAGTGTCTTGTGT	720
OY	721	CTTAATAGTGTCCCATGAGCCAAAGGCTTAATTCACAAAGCCATATGAGAGTGGGGTGTG	780
Db	721	CTTAATAGTGTCCCATGAGCCAAAGGCTTAATTCACAAAGCCATATGAGAGTGGGGTGTG	780
OY	781	GCCATCATCTCTTACCTGAGGCCCCATGATTAATGAGAAAGTGAAGCACTGCAAGTGTGT	840
Db	781	GCCATCATCTCTTACCTGAGGCCCCATGATTAATGAGAAAGTGAAGCACTGCAAGTGTGT	840
OY	841	GCACTTTCTGTGGTAACTATGCTGACACTCTGAGGCCCTGTGAGGTGCTGTGAGCA	900
Db	841	GCACTTTCTGTGGTAACTATGCTGACACTCTGAGGCCCTGTGAGGTGCTGTGAGCA	900
OY	901	AAACCTTCCAAAGAGCTGTGACCTGACGCGCAAGAAACAAAGTCTTTTCACTGAGTGTGT	960
Db	901	AAACCTTCCAAAGAGCTGTGACCTGACGCGCAAGAAACAAAGTCTTTTCACTGAGTGTGT	960
OY	961	GATGTGTCTTTCTTCTAATAGAGCTCTAATGATTTATTTGTCTCAGAAAGCATTTAAACA	1020
Db	961	GATGTGTCTTTCTTCTAATAGAGCTCTAATGATTTATTTGTCTCAGAAAGCATTTAAACA	1020
OY	1021	ATTCTTTCATATGAGAGTCAATTAACAAGAGTGTGCTTCTGTGCTTATGAAGAG	1080
Db	1021	ATTCTTTCATATGAGAGTCAATTAACAAGAGTGTGCTTCTGTGCTTATGAAGAG	1080

Dp	1021	TTTCCTTCACATCCGAGTGCATTAACACAGAGTGTGGCTTCGTGCTCATTAAGAGG	1080
Oy	1081	GCTCCTGAGATCCTCATGTGGGTCCACAAGTCCCTTGCCCTCCATCTGATACAAACATC	1140
Dp	1081	GCTCCTGAGATCCTCAATGTGGGTCCACAAGTCCCTTGCCCTCCATCTGATACAAACATC	1140
Oy	1141	CTGCACATCCCGGCTCAGATATTGTCACCTTGCGCTAATGAATACTTCCATGACAGAC	1200
Dp	1141	CTGCACATCCCGGCTCAGATATTGTCACCTTGCGCTAATGAATACTTCCATGACAGAC	1200
Oy	1201	TCCCTGACTGAATCCGAGACAGTCTTCTGGACTTGTGGAGATGTATTCTTTGTGATC	1260
Dp	1201	TCCCTGACTGAATCCGAGACAGTCTTCTGGACTTGTGGAGATGTATTCTTTGTGATC	1260
Oy	1261	CCTGCACTGATCACAAGCTCGATATCACAAGATGCTGTGTGACACTGTCTACTTATAG	1320
Dp	1261	CCTCACACTGATCACAAGCTCGATATCACAAGATGCTGTGTGACACTGTCTACTTATAG	1320
Oy	1321	TTTGGGACCGGGCTCAGTGTCTTGAAGACACGAAGCCGGCTTTGTCAAGCCGACAC	1380
Dp	1321	TTTGGGACCGGGCTCAGTGTCTTGAAGACACGAAGCCGGCTTTGTCAAGCCGACAC	1380
Oy	1381	GCTGATGAAGTCCGCTTTGTGTGGTGTGCTTCTCTGAAGGGGGACATTGTATATTC	1440
Dp	1381	GCTGATGAAGTCCGCTTTGTGTGGTGTGCTTCTCTGAAGGGGGACATTGTATATTC	1440
Oy	1441	GAAGAGCCACGAGAGAGAGAACTTACTGAGCCGAAGATGATGAATACTGSGGCTACC	1500
Dp	1441	GAAGAGCCACGAGAGAGAGAACTTACTGAGCCGAAGATGATGAATACTGSGGCTACC	1500
Oy	1501	TTTGTCTGGAACCGGGAAATCTTAATGGAGACACTGTCTCTGTGGCCAGCTTAATATCTG	1560
Dp	1501	TTTGTCTGGAACCGGGAAATCTTAATGGAGACACTGTCTCTGTGGCCAGCTTAATATCTG	1560
Oy	1561	ACTGAGCAGTACTCCACAGCTGACCTTGAACATGAGCCTCGGACAGAGACTCAAGAAACCG	1620
Dp	1561	ACTGAGCAGTACTCCACAGCTGACCTTGAACATGAGCCTCGGACAGAGACTCAAGAAACCG	1620
Oy	1621	CGGGTGAATTTTGGACACAGACCAATCCCTGTATCTGTCTGTGCTCGACATGCTCAC	1680
Dp	1621	CGGGTGAATTTTGGACACAGACCAATCCCTGTATCTGTGTGCTCGACATGCTCAC	1680
Oy	1681	AGTCCCTCTTTCCTTAACCTTTCCTCTCTCTCTCCAGGCTTCTTTTCTTTTGTGCT	1740
Dp	1681	AGTCCCTCTTTCCTTAACCTTTCCTCTCTCTCTCTCCAGGCTTCTTTTCTTTTGTGCT	1740
Oy	1741	CCCTGA 1746	
Dp	1741	CCTTGA 1746	

```

? RESULT 3
? US-10-023-515-1
? Sequence 1, Application US/10023515
? Publication No. US2002018636A1
? GENERAL INFORMATION:
? APPLICANT: Curtis, Rory A. J.
? TITLE OF INVENTION: 5310.0, A NOVEL HUMAN CARBOXYESTERASE
? TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
? FILE REFERENCE: 10448-122001
? CURRENT APPLICATION NUMBER: US/10/023,515
? CURRENT FILING DATE: 2001-12-18
? PRIOR APPLICATION NUMBER: 60/356,369
? PRIOR FILING DATE: 2000-12-18
? PRIOR APPLICATION NUMBER: 60/279,508
? PRIOR FILING DATE: 2001-03-28
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 2158
? TYPE: DNA
? ORGANISM: Homo sapiens

```

FEATURE:
; NAME/KEY: CDS
; LOCATION: (96) ... (1838)
US-10-023-515-1

Query Match 100.0%; Score 1746; DB 13; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCCACAGGAGCTTACTTATCTGCTTCAACAGAGTGTCTTTTCTGATTTCTCAGCCC 60
Db 96 ATGCCACAGGAGCTTACTTATCTGCTTCAACAGAGTGTCTTTTCTGATTTCTCAGCCC 60
QY 61 CTGTGGGACACAGACAGTGGGAAAAAAGTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
Db 156 CTGTGGGACACAGACAGTGGGAAAAAAGTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
QY 121 ACCAGGCTGGAGTGGATTCAGGGGACAGCACTGCTGCTGGAGAGCCCTGTGCTGTG 180
Db 216 ACCAGGCTGGAGTGGATTCAGGGGACAGCACTGCTGCTGGAGAGCCCTGTGCTGTG 275
QY 181 AAGGTGTCTCGAGTCCCTTTGTGCTGCCCGCTGGAGTCCCTGGAGATTAGAAC 240
Db 276 AAGGTGTCTCGAGTCCCTTTGTGCTGCCCGCTGGAGTCCCTGGAGATTAGAAC 335
QY 241 CCGCAGCTTGCATGCGCTGGATTAACCTTGGAGAACCACTCTACCTTAATTTGTGC 300
Db 336 CCGCAGCTTGCATGCGCTGGATTAACCTTGGAGAACCACTCTACCTTAATTTGTGC 395
QY 301 CTCACAGAACTCAGAGTGGCTGCTCTTAGATCAACATGCTCAAGGTGATTAACCGAAA 360
Db 396 CTCACAGAACTCAGAGTGGCTGCTCTTAGATCAACATGCTCAAGGTGATTAACCGAAA 455
QY 361 TTGCGAGTGTCAAGAACTGCTCTTACCTGAACATGCTGCTGCCCTGACCCGATACA 420
Db 456 TTGCGAGTGTCAAGAACTGCTCTTACCTGAACATGCTGCTGCCCTGACCCGATACA 515
QY 421 GGGTCAAGCTCCCGCTTTGGTGTGGTCCCAAGAGTGGCTTCAAGCTGCTCAACC 480
Db 516 GGGTCAAGCTCCCGCTTTGGTGTGGTCCCAAGAGTGGCTTCAAGCTGCTCAACC 575
QY 481 TCCATCTTTGATGGGTCCGCTGGCTGCTGATGAGAGAGCTGGTGTGGTGTGCTCAG 540
Db 576 TCCATCTTTGATGGGTCCGCTGGCTGCTGATGAGAGAGCTGGTGTGGTGTGCTCAG 635
QY 541 TACCGGCTAGAGAAATTTGGTCTTCAACACATGAGTCAAGCATGCTCCGAGAACTGG 600
Db 636 TACCGGCTAGAGAAATTTGGTCTTCAACACATGAGTCAAGCATGCTCCGAGAACTGG 695
QY 601 GCCTTCAAGAGACAGAGTGGCTGCTGATGAGAGAGTGGTGGTGGTGGTGGTGGTGG 660
Db 696 GCCTTCAAGAGACAGAGTGGCTGCTGATGAGAGAGTGGTGGTGGTGGTGGTGGTGG 755
QY 661 GGGGACCCCAAGCTGTGACCATTTTGGAGTCCGCGGAGGCATTAAGTGTCTAGT 720
Db 756 GGGGACCCCAAGCTGTGACCATTTTGGAGTCCGCGGAGGCATTAAGTGTCTAGT 815
QY 721 CTTAATCTGTCTCCCATGAGCCCAAGGCTTATTCACAAAGCATCAGAGAGTGGGCTG 780
Db 816 CTTAATCTGTCTCCCATGAGCCCAAGGCTTATTCACAAAGCATCAGAGAGTGGGCTG 875
QY 781 GCCATCATCTCTTACCTGAGAGGCCATGATTATAGAGAGTGGAGACTGCGAGTGGT 840
Db 876 GCCATCATCTCTTACCTGAGAGGCCATGATTATAGAGAGTGGAGACTGCGAGTGGT 935
QY 841 GGCATTTTGTGTGAACATGCTCAGACTCTGAGGCTTGTGAGTGGCTGAGAGACA 900
Db 936 GGCATTTTGTGTGAACATGCTCAGACTCTGAGGCTTGTGAGTGGCTGAGAGACA 995
QY 901 AAACCTCTCAAGAGCTGCTGACCTCAGGCAAGAAACAAAGCTTTCATCTCAGAGTGT 960
Db 996 AAACCTCTCAAGAGCTGCTGACCTCAGGCAAGAAACAAAGCTTTCATCTCAGAGTGT 1055
```

```
QY 961 GATGTGCTTTCTTCTTATGAGCCTTAGATCTATTGTCTCAGAAAGATTTAAAGCA 1020
Db 1056 GATGTGCTTTCTTCTTATGAGCCTTAGATCTATTGTCTCAGAAAGATTTAAAGCA 1115
QY 1021 ATTCCCTCATCATGAGAGTCAATAACCAAGTGTGCTTCTGCTCTATGAGAG 1080
Db 1116 ATTCCCTCATCATGAGAGTCAATAACCAAGTGTGCTTCTGCTCTATGAGAG 1175
QY 1081 GCTCTGAGATCTCAGTGTCTCAACAAGTCCCTTGGCCCTCATCTGATCAAAACATC 1140
Db 1176 GCTCTGAGATCTCAGTGTCTCAACAAGTCCCTTGGCCCTCATCTGATCAAAACATC 1235
QY 1141 CTGCAATCTCCGCTCAGTATTGCACTTGTGGCTAATGAATCTTCCATGACAGAC 1200
Db 1236 CTGCAATCTCCGCTCAGTATTGCACTTGTGGCTAATGAATCTTCCATGACAGAC 1295
QY 1201 TCCCTGATGAAATCCGACAGAGTCTTGTGACCTTGTGAGATGTGTTTGTGTGCTC 1260
Db 1296 TCCCTGATGAAATCCGACAGAGTCTTGTGAGATGTGTTTGTGTGCTC 1355
QY 1261 CCTGATCTGATCAGAGTCAATATCAGAGAGTCTGATGACCTGCTATCTATGAG 1320
Db 1356 CCTGATCTGATCAGAGTCAATATCAGAGAGTCTGATGACCTGCTATCTATGAG 1415
QY 1321 TTGCGGACCGGCTCAGTGTGTTGAAGACAGAGCCGCTTTGTCAAGCCGACAC 1380
Db 1416 TTGCGGACCGGCTCAGTGTGTTGAAGACAGAGCCGCTTTGTCAAGCCGACAC 1475
QY 1381 GCTGATGAGTCCGCTTGTGTTGCTGATGAGTCCCTTCTGAAGGGGACATGTTATGTC 1440
Db 1476 GCTGATGAGTCCGCTTGTGTTGCTGATGAGTCCCTTCTGAAGGGGACATGTTATGTC 1535
QY 1441 GAAGAGCCACAGAGAGGAGAAATTAAGAGCCGAGAGTGAATTAAGAGCTAC 1500
Db 1536 GAAGAGCCACAGAGAGGAGAAATTAAGAGCCGAGAGTGAATTAAGAGCTAC 1595
QY 1501 TTGCTGAAACCGGAAATCTTAATGAGAAACAGCTGCTCTGAGGCACTTATATCTG 1560
Db 1596 TTGCTGAAACCGGAAATCTTAATGAGAAACAGCTGCTCTGAGGCACTTATATCTG 1655
QY 1561 ACTGAGAGTACTCTCAAGCTGACCTTGAACATGAGGCTCGGACAGAGCTCAAGAACCG 1620
Db 1656 ACTGAGAGTACTCTCAAGCTGACCTTGAACATGAGGCTCGGACAGAGCTCAAGAACCG 1715
QY 1621 CGGAGTGAATTTTGGACCAAGCAATCCCTGATCTGTGCTCCGACATGCTCCAC 1680
Db 1716 CGGAGTGAATTTTGGACCAAGCAATCCCTGATCTGTGCTCCGACATGCTCCAC 1775
QY 1681 AGTCTCTTCTCTTAATCTTCTCTCTCTCTCAGCTTCTTCTTCTTGTGTGT 1740
Db 1776 AGTCTCTTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
QY 1741 CCTTGA 1746
Db 1836 CCTTGA 1841
```

RESULT 4
US-10-674-636-1
; Sequence 1, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674, 636
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: 60/279,508
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2158
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (96)...(1838)
 ; US-10-674-636-1

Query Match 100.0%; Score 1746; DB 18; Length 2158;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCAGAGGACTTACTTCAATCTCTCAATGAGTCTTTCTGATTTCTCAGGCC 60
 DB ATGCCAGAGGACTTACTTCAATCTCTCAATGAGTCTTTCTGATTTCTCAGGCC 155
 QY 61 CTGTTGGGACACAGCAGTGGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
 DB CTGTTGGGACACAGCAGTGGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
 QY 121 ACCAGGCTGGATGATTCAGGGGCAAGCAATCATGCTGGGAAGCCCTGGCTGTG 180
 DB ACCAGGCTGGATGATTCAGGGGCAAGCAATCATGCTGGGAAGCCCTGGCTGTG 275
 QY 216 ACCAGGCTGGATGATTCAGGGGCAAGCAATCATGCTGGGAAGCCCTGGCTGTG 275
 DB ACCAGGCTGGATGATTCAGGGGCAAGCAATCATGCTGGGAAGCCCTGGCTGTG 275
 QY 181 AACGTTCTCTGGAGTCCCTTTGCTGCTGCCGCTGGGATCCCTGCAATTTACGAA 240
 DB AACGTTCTCTGGAGTCCCTTTGCTGCTGCCGCTGGGATCCCTGCAATTTACGAA 335
 QY 276 AACGTTCTCTGGAGTCCCTTTGCTGCTGCCGCTGGGATCCCTGCAATTTACGAA 335
 DB AACGTTCTCTGGAGTCCCTTTGCTGCTGCCGCTGGGATCCCTGCAATTTACGAA 335
 QY 241 CCGCAGCTGCATCCGCTTGGGATTAATCTGGCAAGGCACTCTCTAATTTTGTG 300
 DB CCGCAGCTGCATCCGCTTGGGATTAATCTGGCAAGGCACTCTCTAATTTTGTG 395
 QY 336 CCGCAGCTGCATCCGCTTGGGATTAATCTGGCAAGGCACTCTCTAATTTTGTG 395
 DB CCGCAGCTGCATCCGCTTGGGATTAATCTGGCAAGGCACTCTCTAATTTTGTG 395
 QY 301 CTCGAGACTGAGATGGCTCTTAATCAACATGCTCAAGGTGATTAACCGGAA 360
 DB CTCGAGACTGAGATGGCTCTTAATCAACATGCTCAAGGTGATTAACCGGAA 455
 QY 396 CTCGAGACTGAGATGGCTCTTAATCAACATGCTCAAGGTGATTAACCGGAA 455
 DB CTCGAGACTGAGATGGCTCTTAATCAACATGCTCAAGGTGATTAACCGGAA 515
 QY 456 TTCGAGGTGAGAGACTGCTCTTAATCAACATGCTCAAGGTGATTAACCGGAA 515
 DB TTCGAGGTGAGAGACTGCTCTTAATCAACATGCTCAAGGTGATTAACCGGAA 515
 QY 421 GGGCTCAAGCTCCCGCTTGGTGTGTTCCAGAGGTGCTTCAAGCTGCTCAGCC 480
 DB GGGCTCAAGCTCCCGCTTGGTGTGTTCCAGAGGTGCTTCAAGCTGCTCAGCC 575
 QY 516 GGGCTCAAGCTCCCGCTTGGTGTGTTCCAGAGGTGCTTCAAGCTGCTCAGCC 575
 DB GGGCTCAAGCTCCCGCTTGGTGTGTTCCAGAGGTGCTTCAAGCTGCTCAGCC 575
 QY 481 TCCATCTTGAATGGTCCGCTGCTGCTAATGAGAGCTGCTGGTGGTCCGCAAG 540
 DB TCCATCTTGAATGGTCCGCTGCTGCTAATGAGAGCTGCTGGTGGTCCGCAAG 635
 QY 576 TCCATCTTGAATGGTCCGCTGCTGCTAATGAGAGCTGCTGGTGGTCCGCAAG 635
 DB TCCATCTTGAATGGTCCGCTGCTGCTAATGAGAGCTGCTGGTGGTCCGCAAG 635
 QY 541 TACCGGCTAGGAAATTTTGTCTTCAACATGAGGATCAGCATGCTCCGGGAACTGG 600
 DB TACCGGCTAGGAAATTTTGTCTTCAACATGAGGATCAGCATGCTCCGGGAACTGG 695
 QY 636 TACCGGCTAGGAAATTTTGTCTTCAACATGAGGATCAGCATGCTCCGGGAACTGG 695
 DB TACCGGCTAGGAAATTTTGTCTTCAACATGAGGATCAGCATGCTCCGGGAACTGG 695
 QY 601 GCGTTCAAGACCAAGTGGCTGCTGCTGCTGGGCTCAGAGAACATGAGTTCTTGGT 660
 DB GCGTTCAAGACCAAGTGGCTGCTGCTGCTGGGCTCAGAGAACATGAGTTCTTGGT 755
 QY 696 GCGTTCAAGACCAAGTGGCTGCTGCTGCTGGGCTCAGAGAACATGAGTTCTTGGT 755
 DB GCGTTCAAGACCAAGTGGCTGCTGCTGCTGGGCTCAGAGAACATGAGTTCTTGGT 755
 QY 661 GGGGAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 DB GGGGAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
 QY 756 GGGGAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
 DB GGGGAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
 QY 721 CTTATACTGTCTCCAGTGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGGTG 780
 DB CTTATACTGTCTCCAGTGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGGTG 875
 QY 816 CTTATACTGTCTCCAGTGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGGTG 875
 DB CTTATACTGTCTCCAGTGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGGTG 875
 QY 781 GGCATCATCTCTTACCTGAGGCGCATGATTTATGAGAGAGTGAAGACCTGCAAGTGTG 840
 DB GGCATCATCTCTTACCTGAGGCGCATGATTTATGAGAGAGTGAAGACCTGCAAGTGTG 935

QY 841 GCACATTTCTGTGTAAATGCGTCAAGCTCTGAGGCCCTGCTGAGTGCCTGAGACA 900
 DB GCACATTTCTGTGTAAATGCGTCAAGCTCTGAGGCCCTGCTGAGTGCCTGAGACA 995
 QY 936 GCACATTTCTGTGTAAATGCGTCAAGCTCTGAGGCCCTGCTGAGTGCCTGAGACA 995
 DB GCACATTTCTGTGTAAATGCGTCAAGCTCTGAGGCCCTGCTGAGTGCCTGAGACA 995
 QY 901 AAGCCCTCAAGAGAGCTGCTGAGCCCTCAGCCAGAAAAAAGATCTTTCACTGAGTGT 960
 DB AAGCCCTCAAGAGAGCTGCTGAGCCCTCAGCCAGAAAAAAGATCTTTCACTGAGTGT 1055
 QY 996 AAGCCCTCAAGAGAGCTGCTGAGCCCTCAGCCAGAAAAAAGATCTTTCACTGAGTGT 1055
 DB AAGCCCTCAAGAGAGCTGCTGAGCCCTCAGCCAGAAAAAAGATCTTTCACTGAGTGT 1055
 QY 961 GATGAGCTTTCTTTCTTAATGAGCCTCAATGATCTTATGCTCAGAAAGATTTAAAGA 1020
 DB GATGAGCTTTCTTTCTTAATGAGCCTCAATGATCTTATGCTCAGAAAGATTTAAAGA 1115
 QY 1056 GATGAGCTTTCTTTCTTAATGAGCCTCAATGATCTTATGCTCAGAAAGATTTAAAGA 1115
 DB GATGAGCTTTCTTTCTTAATGAGCCTCAATGATCTTATGCTCAGAAAGATTTAAAGA 1115
 QY 1021 ATTCTTCATCATGAGTCAATTAACAGAGTGGCTTCTGCTGCTATGAAGAG 1080
 DB ATTCTTCATCATGAGTCAATTAACAGAGTGGCTTCTGCTGCTATGAAGAG 1175
 QY 1116 ATTCTTCATCATGAGTCAATTAACAGAGTGGCTTCTGCTGCTATGAAGAG 1175
 DB ATTCTTCATCATGAGTCAATTAACAGAGTGGCTTCTGCTGCTATGAAGAG 1175
 QY 1081 GCTCTGAGATCTCAAGTGGCTCAGCAAGTCCCTTGCCTTCATCTGATTAACAATC 1140
 DB GCTCTGAGATCTCAAGTGGCTCAGCAAGTCCCTTGCCTTCATCTGATTAACAATC 1235
 QY 1176 GCTCTGAGATCTCAAGTGGCTCAGCAAGTCCCTTGCCTTCATCTGATTAACAATC 1235
 DB GCTCTGAGATCTCAAGTGGCTCAGCAAGTCCCTTGCCTTCATCTGATTAACAATC 1235
 QY 1141 CTGCAATCCCGGCTCAGTATTTGACCTTGGCTAATGAATATTCATGACAGAGAC 1200
 DB CTGCAATCCCGGCTCAGTATTTGACCTTGGCTAATGAATATTCATGACAGAGAC 1295
 QY 1236 CTGCAATCCCGGCTCAGTATTTGACCTTGGCTAATGAATATTCATGACAGAGAC 1295
 DB CTGCAATCCCGGCTCAGTATTTGACCTTGGCTAATGAATATTCATGACAGAGAC 1355
 QY 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGCACTTCTGAGATGTCTTTGTGCTC 1260
 DB TCCCTGACTGAATCCGAGACAGTCTTCTGCACTTCTGAGATGTCTTTGTGCTC 1355
 QY 1261 CCTGCACTGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTATATAG 1320
 DB CCTGCACTGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTATATAG 1415
 QY 1356 CCTGCACTGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTATATAG 1415
 DB CCTGCACTGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTATATAG 1415
 QY 1321 TTTGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTGTGAAGCGAGCAC 1380
 DB TTTGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTGTGAAGCGAGCAC 1475
 QY 1416 TTTGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTGTGAAGCGAGCAC 1475
 DB TTTGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTGTGAAGCGAGCAC 1535
 QY 1381 GCTGATGAAGTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGGAATTTATGCTC 1440
 DB GCTGATGAAGTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGGAATTTATGCTC 1535
 QY 1476 GCTGATGAAGTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGGAATTTATGCTC 1535
 DB GCTGATGAAGTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGGAATTTATGCTC 1535
 QY 1441 GAAAGGACCAAGAGAGAGAGAGTAACTGAGCCGGAAGATGATGAATATCTGGCTAC 1500
 DB GAAAGGACCAAGAGAGAGAGAGTAACTGAGCCGGAAGATGATGAATATCTGGCTAC 1595
 QY 1536 GAAAGGACCAAGAGAGAGAGAGTAACTGAGCCGGAAGATGATGAATATCTGGCTAC 1595
 DB GAAAGGACCAAGAGAGAGAGAGTAACTGAGCCGGAAGATGATGAATATCTGGCTAC 1595
 QY 1501 TTTGCTGAAACCGGGAATCTTAATGGGAAAGCACTGCTGCTGCTGCTGCTGCTGCT 1560
 DB TTTGCTGAAACCGGGAATCTTAATGGGAAAGCACTGCTGCTGCTGCTGCTGCTGCT 1655
 QY 1596 TTTGCTGAAACCGGGAATCTTAATGGGAAAGCACTGCTGCTGCTGCTGCTGCTGCT 1655
 DB TTTGCTGAAACCGGGAATCTTAATGGGAAAGCACTGCTGCTGCTGCTGCTGCTGCT 1655
 QY 1561 ACTGAGAGTACCTTCAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1620
 DB ACTGAGAGTACCTTCAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1715
 QY 1656 ACTGAGAGTACCTTCAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1715
 DB ACTGAGAGTACCTTCAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1715
 QY 1621 CGGGTGAATTTTGGACAGAGACATCCCTGATCTGTCTGCTCGACATGCTCCAG 1680
 DB CGGGTGAATTTTGGACAGAGACATCCCTGATCTGTCTGCTCGACATGCTCCAG 1775
 QY 1716 CGGGTGAATTTTGGACAGAGACATCCCTGATCTGTCTGCTCGACATGCTCCAG 1775
 DB CGGGTGAATTTTGGACAGAGACATCCCTGATCTGTCTGCTCGACATGCTCCAG 1775
 QY 1681 AGTCTCTTTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
 DB AGTCTCTTTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
 QY 1776 AGTCTCTTTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
 DB AGTCTCTTTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
 QY 1741 CTTGA 1746
 DB CTTGA 1841

RESULT 5
 US-10-757-262-45
 ; Sequence 45, Application US/10757262
 ; Publication No. US20040197825A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Karichehli, Venkateswarlu
 ; APPLICANT: Siles-Santiago, Immaculada
 ; APPLICANT: Eliasof, Scott D.

```

/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44350, 54181, 211, 5687, 884,
/ TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
/ TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43373, 51164,
/ TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
/ TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
/ TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
/ TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
/ TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
/ FILE REFERENCE: MP103-007P1RNONMIN
/ CURRENT APPLICATION NUMBER: US/10/757, 262
/ CURRENT FILING DATE: 2004-01-14
/ PRIOR APPLICATION NUMBER: US 60/440,318
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/444,783
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: US 60/457,901
/ PRIOR FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: US 60/468,775
/ PRIOR FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/471,614
/ PRIOR FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: US 60/478,742
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/488,529
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/491,156
/ PRIOR FILING DATE: 2003-07-30
/ PRIOR APPLICATION NUMBER: US 60/499,594
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 60/506,332
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 45
/ LENGTH: 2158
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (96) ... (1841)
/ US-10-757-262-45

Query Match      100.0%; Score 1746; DB 19; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCCACAGGACTTACTTCATCTGCTTCACAAATGCTGCTTTTCTCGATTCCTCAGCCC 60
DB      96 ATGCCACAGGACTTACTTCATCTGCTTCACAAATGCTGCTTTTCTCGATTCCTCAGCCC 155
QY      61 CTGTTGGGACACACAGCATGTGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAAAC 120
DB      156 CTGTTGGGACACACAGCATGTGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAAAC 215
QY      121 ACCAGGCTGGATGATTCAGGGACAGCACTGTGCTGGGAACCCCTGTGCTGTG 180
DB      216 ACCAGGCTGGATGATTCAGGGACAGCACTGTGCTGGGAACCCCTGTGCTGTG 275
QY      181 AACCTGTTCTCGAGATCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTAGCAAC 240
DB      276 AACCTGTTCTCGAGATCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTAGCAAC 335
QY      241 CCGGAGCTGCTGATGCTGGGATTAATGCGAGAAACCACTTCATACCTTAATTTGTGC 300
DB      336 CCGGAGCTGCTGATGCTGGGATTAATGCGAGAAACCACTTCATACCTTAATTTGTGC 395
QY      301 CTCGAAATCTCAGAGTGTGCTTTAGATCAACATGCTCAAGGTGATTAACCGGAAA 360
DB      396 CTCGAAATCTCAGAGTGTGCTTTAGATCAACATGCTCAAGGTGATTAACCGGAAA 455
QY      361 TTGGAAGTGTCAAGAGCTGCTCTTACCTGAACATCTATGCGCTGCCACCGCGATACA 420
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44350, 54181, 211, 5687, 884,
/ TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
/ TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43373, 51164,
/ TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
/ TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
/ TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
/ TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
/ TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
/ FILE REFERENCE: MP103-007P1RNONMIN
/ CURRENT APPLICATION NUMBER: US/10/757, 262
/ CURRENT FILING DATE: 2004-01-14
/ PRIOR APPLICATION NUMBER: US 60/440,318
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/444,783
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: US 60/457,901
/ PRIOR FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: US 60/468,775
/ PRIOR FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/471,614
/ PRIOR FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: US 60/478,742
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/488,529
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/491,156
/ PRIOR FILING DATE: 2003-07-30
/ PRIOR APPLICATION NUMBER: US 60/499,594
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 60/506,332
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 45
/ LENGTH: 2158
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (96) ... (1841)
/ US-10-757-262-45

Query Match      100.0%; Score 1746; DB 19; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCCACAGGACTTACTTCATCTGCTTCACAAATGCTGCTTTTCTCGATTCCTCAGCCC 60
DB      96 ATGCCACAGGACTTACTTCATCTGCTTCACAAATGCTGCTTTTCTCGATTCCTCAGCCC 155
QY      61 CTGTTGGGACACACAGCATGTGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAAAC 120
DB      156 CTGTTGGGACACACAGCATGTGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAAAC 215
QY      121 ACCAGGCTGGATGATTCAGGGACAGCACTGTGCTGGGAACCCCTGTGCTGTG 180
DB      216 ACCAGGCTGGATGATTCAGGGACAGCACTGTGCTGGGAACCCCTGTGCTGTG 275
QY      181 AACCTGTTCTCGAGATCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTAGCAAC 240
DB      276 AACCTGTTCTCGAGATCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTAGCAAC 335
QY      241 CCGGAGCTGCTGATGCTGGGATTAATGCGAGAAACCACTTCATACCTTAATTTGTGC 300
DB      336 CCGGAGCTGCTGATGCTGGGATTAATGCGAGAAACCACTTCATACCTTAATTTGTGC 395
QY      301 CTCGAAATCTCAGAGTGTGCTTTAGATCAACATGCTCAAGGTGATTAACCGGAAA 360
DB      396 CTCGAAATCTCAGAGTGTGCTTTAGATCAACATGCTCAAGGTGATTAACCGGAAA 455
QY      361 TTGGAAGTGTCAAGAGCTGCTCTTACCTGAACATCTATGCGCTGCCACCGCGATACA 420
```

Db 1536 GAAGGAGCCAGGAGGAGAGATTACTGAGCCGAGAGATGATGAATACTAGGAGTACC 1595
Qy 1501 TTGTGCGAAGCCGGGAATCTATATGGGAAGACCTGTCTGTGGCCAGCTTATATCTG 1560
Db 1536 TTGTGCGAAGCCGGGAATCTATATGGGAAGACCTGTCTGTGGCCAGCTTATATCTG 1555
Qy 1561 ACTGAGCAGTACCTCCAGCTGGAATTGAAATGAGGCTTGGACAGAGACTCAAGAACCG 1620
Db 1656 ACTGAGCAGTACCTCCAGCTGGAATTGAAATGAGGCTTGGACAGAGACTCAAGAACCG 1715
Qy 1621 CGGGTGAATTTTGGACCAAGCAACATCCCCCTGATCTGTCTGCTCCGACATGCTCCAC 1680
Db 1716 CGGGTGAATTTTGGACCAAGCAACATCCCCCTGATCTGTCTGCTCCGACATGCTCCAC 1775
Qy 1681 AGTCCCTTTCTTCTTAACTTTCCTCTCTCTCCAGCCTTTCTTTCTTTTGTGCT 1740
Db 1776 AGTCCCTTTCTTCTTAACTTTCCTCTCTCTCTCCAGCCTTTCTTTCTTTTGTGCT 1835
Qy 1741 CCTTGA 1746
Db 1836 CCTTGA 1841

RESULT 6

US-10-451-168-40
; Sequence 40, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GPs0039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-40

Query Match 94.7%; Score 1652.8; DB 18; Length 1728;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 91 GGGCTTGTGTGAAGGGGCAAGAGAAACACAGGCTGGATGATTCAGGGCAAGCA 150
Db 73 GGGCTTGTGTGAAGGGGCAAGAGAAACACAGGCTGGATGATTCAGGGCAAGCA 132
Qy 151 GTGCTGTGTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGAGAGTCCCTTTGTGCT 210
Db 133 GTGCTGTGTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGAGAGTCCCTTTGTGCT 192

Qy 211 CCCCGCTGGATCCCTGGATTATAGAACCCGAGCCTGATCGCCCTGGATTAATTG 270
Db 199 CCCCGCTGGATCCCTGGATTATAGAACCCGAGCCTGATCGCCCTGGATTAATTG 252
Qy 271 CGAAGAGCACTCTTACCTTAATTTTGTGCTCCAGAACTCAGAGTGGTGTCTTAAGT 330
Db 253 CGAAGAGCACTCTTACCTTAATTTTGTGCTCCAGAACTCAGAGTGGTGTCTTAAGT 312
Qy 331 CAACATGCTCAAGGTGATTAACCGAAATTGGAGTGTGAGAAAGCTGCTTAACCTG 390
Db 313 CAACATGCTCAAGGTGATTAACCGAAATTGGAGTGTGAGAAAGCTGCTTAACCTG 372
Qy 391 AACATCTATGCGCTGCCACCGCATACAGGCTCCAACTCCCGCTTTGTGTGCTTC 450
Db 373 AACATCTATGCGCTGCCACCGCATACAGGCTCCAACTCCCGCTTTGTGTGCTTC 432
Qy 451 CCAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTTGATGGGTCCGCTGTGCTC 510
Db 433 CCAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTTGATGGGTCCGCTGTGCTC 492
Qy 511 TATGAGAGCGTGTGTTGTGTGTGCTTCAAGTACCGGCTAGGAATATTGTTCTTACCC 570
Db 493 TATGAGAGCGTGTGTTGTGTGTGCTTCAAGTACCGGCTAGGAATATTGTTCTTACCC 552
Qy 571 ACATGGGATCAGCATGCTCCGCGGAATCGGCTTCAAGACCAAGTGTGCTGTCTC 630
Db 553 ACATGGGATCAGCATGCTCCGCGGAATCGGCTTCAAGACCAAGTGTGCTGTCTC 612
Qy 631 TGGGTCCAGAGACATGAGTGTCTTGGTGGGAACCCAGCTGTGTGACATCTTTGGC 690
Db 613 TGGGTCCAGAGACATGAGTGTCTTGGTGGGAACCCAGCTGTGTGACATCTTTGGC 672
Qy 691 GAGTCCGCGGAGCATAGTGTCTTCAAGTGTCTTCAAGTGTCTTCAAGTGTCTTCAAG 750
Db 673 GAGTCCGCGGAGCATAGTGTCTTCAAGTGTCTTCAAGTGTCTTCAAGTGTCTTCAAG 732
Qy 751 TTCACAAAGCCATGAGAGAGTGGGTGGCCATCATCTCTTACCTGAGGCCATGAT 810
Db 733 TTCACAAAGCCATGAGAGAGTGGGTGGCCATCATCTCTTACCTGAGGCCATGAT 792
Qy 811 TATGAGAGAGTGAAGACCTGAGGTGTGCAATTTCTGTGTGAACATGCTGACAG 870
Db 793 TATGAGAGAGTGAAGACCTGAGGTGTGCAATTTCTGTGTGAACATGCTGACAG 852
Qy 871 TCTGAGGCGCTGCTGAGGTGTGCAAAACCTCCAAAGAGCTGTGACCTTCAGC 930
Db 853 TCTGAGGCGCTGCTGAGGTGTGCAAAACCTCCAAAGAGCTGTGACCTTCAGC 912
Qy 931 CAGAAACAAAGTCTTTCATCTGAGGTGTGATGTGCTTTCTTCTTAATAGCCTCTA 990
Db 913 CAGAAACAAAGTCTTTCATCTGAGGTGTGATGTGCTTTCTTCTTAATAGCCTCTA 972
Qy 991 GATCTATTGTCTCAGAAAGCATTTAAGCAATCTTCCATCATGAGTCAATTAACAC 1050
Db 973 GATCTATTGTCTCAGAAAGCATTTAAGCAATCTTCCATCATGAGTCAATTAACAC 1032
Qy 1051 GAGTGTGCTTCTGCTGCTTATGAGAGAGTCTCTGAGATCTTCAAGTGTGCTTCAAG 1110
Db 1033 GAGTGTGCTTCTGCTGCTTATGAGAGAGTCTCTGAGATCTTCAAGTGTGCTTCAAG 1092
Qy 1111 TCCCTTGGCTTCATCTGATACAAACATCTGCACTCCGCTCAGTATTTGACCTT 1170
Db 1093 TCCCTTGGCTTCATCTGATACAAACATCTGCACTCCGCTCAGTATTTGACCTT 1152
Qy 1171 GTGGCTAATGATTAATCTTCATGACAAAGCATCTCCGCTGAGTGAATCCAGAGCTTCTG 1230
Db 1153 GTGGCTAATGATTAATCTTCATGACAAAGCATCTCCGCTGAGTGAATCCAGAGCTTCTG 1212
Qy 1231 GACTTGTGTGAGATGATGTTCTTTGTGTGCTTCCCTGACATGATCAAGCTGATATACAGA 1290
Db 1213 GACTTGTGTGAGATGATGTTCTTTGTGTGCTTCCCTGACATGATCAAGCTGATATACAGA 1272
Qy 1291 GATGCTGTGACACTGTCTACTTCTATGAGTTTGGGACCGGCTCAGTGTCTTGAAGAC 1350

Db 1273 GATGCTGGTGCACTGCTCTACTTCTATGATGTTTCGGACACCGGCTTCAGTGGCTTTGGAAAGAC 13322

QY 1351 ACGAAGCCGGCTTTTGTCTAAAGCCGACCAACGCTGATGTAAGTCCGCTTTGTGTTGGCGGT 1410

Db 1333 ACGAAGCCAGCTTTTGTCTAAAGCCGACCAACGCTGATGTAAGTCCGCTTTGTGTTGGCGGT 13922

QY 1411 GCCTTCTCGAAGGGGGGACATTGTTATGTTTCGAAGGACCGAGAGAGAAAGTTACTG 1470

Db 1393 GCCTTCTCGAAGGGGGGACATTGTTATGTTTCGAAGGACCGAGAGAGAAAGTTACTG 1452

QY 1471 AGCCGGAAGATGATGAATAATCTGGGCTACCTTTGCTCGAACCGGGAAATCTTAATGGGAAC 1530

Db 1453 AGCCGGAAGATGATGAATAATCTGGGCTACCTTTGCTCGAACCGGGAAATCTTAATGGGAAC 1512

QY 1511 GACCTGTCTCTGGGCGCAGCTTATATATCTGACTGAGCACTACCTTCAGCTGGACTTGAAC 1590

Db 1513 GACCTGTCTCTGGGCGCAGCTTATATATCTGACTGAGCACTACCTTCAGCTGGACTTGAAC 1572

QY 1591 ATGAGCCTCGGACAGAGACTCAAGAAACCGGGGTGATTTTGGACGAGACCATGCC 1650

Db 1573 ATGAGCCTCGGACAGAGACTCAAGAAACCGGGGTGATTTTGGACGAGACCATGCC 1632

QY 1651 CTGATCTCTGTCTGGCTCCGACATCTCCACAGTCTCTTTCTTCTTAATCTTCTCTCT 1710

Db 1633 CTGATCTCTGTCTGGCTCCGACATCTCCACAGTCTCTCTTCTTCTTAATCTTCTCTCT 1692

QY 1711 CTCTCTCAGGCTTCTTTTCTTTTGTGCTCTTGA 1746

Db 1693 CTCTCTCAGGCTTCTTTTCTTTTGTGCTCTTGA 1728

```

RESULT 7
US-10-433--256--23
; Sequence 23, Application US/10433256
; Publication No. US20040081980A1
; GENERAL INFORMATION:
; APPLICANT: SANJANWALA, Madhusudan M.; YAO, Monique G.
; APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
; APPLICANT:
; APPLICANT: ARVIZU, Chandra S.; RING, Huijun Z.
; APPLICANT: LEE, Ernestine A.; DING, Li
; APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
; APPLICANT: YUE, Henry; TRIBOUTLEY, Catherine M.
; APPLICANT: LU, Dying; ALINA M.; LAL, Preeti G.
; APPLICANT: WARREN, Bridget A.; YANG, Junming
; APPLICANT: CHAWLA, Narinder K.; NGUYEN, Damien B.
; APPLICANT: GANDHI, Ameena R.; LU, Yan
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0313 USN
; CURRENT APPLICATION NUMBER: US/10/433,256
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/47429
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/254,308
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/256,189
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/257,713
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/262,706
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/266,020
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 23
; LENGTH: 2232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 6538080CB1

```

US-10-433-256-23

Query Match	94.6%	Score 1651.6;	DB 18;	Length 2232;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1657; Conservative	0;	Mismatches	9;	Indels 0; Gaps 0

Qy	81	GGGAAAAA	CTGGG	CTTTCG	TGAAAGG	GCACAGAGAA	CACCGAG	CTGGATG	GAATTTCA	140	
Db	264	GGTTC	CAAGAGG	GCCTTCTG	CTGAAGGG	CCACAGAGAA	CAC	C	AGGCTGGATG	GAATTTCA	323
Qy	141	GGGCAAGCA	AGTCACTG	TCG	TGGAGAGCC	CTGTG	CTGGAACGTG	TTCTCTG	CGAGTCC	200	
Db	324	GGGCAAGCA	AGTCACTG	CTG	TGGAGAGCC	CTGTG	CTGGAACGTG	TTCTCTG	CGAGTCC	383	
Qy	201	CTTTG	CTGCTCC	CCGCTG	GGGATCC	CTGGATTTACGA	CCCGCAG	CCCTGCAT	TCGCTG	260	
Db	384	CTTTG	CTGCTCC	CCGCTG	GGGATCC	CTGGATTTACGA	CCCGCAG	CCCTGCAT	TCGCTG	443	
Qy	261	GGATTA	CTTGCAG	AAGCAC	CTCTCA	CTTAATTG	AGCTCCGAA	CTCAGAGT	GGCT	320	
Db	444	GGATTA	CTTGCAG	AAGCAC	CTCTCA	CTTAATTG	AGCTCCGAA	CTCAGAGT	GGCT	503	
Qy	321	GGCTTTA	GTAAACA	CATG	CTCAAGST	GCATTTACCGAA	TTCCGAGT	GCAGAA	CTG	380	
Db	504	GGCTTTA	GTAAACA	CATG	CTCAAGST	GCATTTACCGAA	TTCCGAGT	GCAGAA	CTG	563	
Qy	381	CCCTTA	CCCTGAACAT	CTATAGG	CCCTGCCACAG	CCGATACAGG	CTCCCGCTT			440	
Db	564	CCCTTA	CCCTGAACAT	CTATAGG	CCCTGCCACAG	CCGATACAGG	CTCCCGCTT			623	
Qy	441	GGTGTGT	TTCCAGAGAG	TGCTTCC	AGACTGG	CTCACTG	CTCATCTTTGAT	TGGGTCCG		500	
Db	624	GGTGTGT	TTCCAGAGAG	TGCTTCC	AGACTGG	CTCACTG	CTCATCTTTGAT	TGGGTCCG		683	
Qy	501	CCCTG	CTCCTATGAGAG	CACTG	CTGTTGTGTGTG	CTGCACAGTA	CCGAGCTAG	GAATATTTG		560	
Db	684	CCCTG	CTCCTATGAGAG	CACTG	CTGTTGTGTGTG	CTGCACAGTA	CCGAGCTAG	GAATATTTG		743	
Qy	561	TTTCTT	CAACA	CATGGGAT	TCAGCATG	CTCCGGGAA	CTGGCCCTT	CAAGCA	CAAGTGGC	620	
Db	744	TTTCTT	CAACA	CATGGGAT	TCAGCATG	CTCCGGGAA	CTGGCCCTT	CAAGCA	CAAGTGGC	803	
Qy	621	TGCTGT	CTGCTGGG	TCAGAA	GAACAT	TCGAGTCTT	CGTGGGGA	CCCGAG	CTGTGAC	680	
Db	804	TGCTGT	CTGCTGGG	TCAGAA	GAACAT	TCGAGTCTT	CGTGGGGA	CCCGAG	CTGTGAC	863	
Qy	681	CATCTT	GGGAG	GTCCCG	GGAGCCAT	AACTGTTCTA	GTCTTAA	CTGTCTCC	CAATGAC	740	
Db	864	CATCTT	GGGAG	GTCCCG	GGAGCCAT	AACTGTTCTA	GTCTTAA	CTGTCTCC	CAATGAC	923	
Qy	741	CAAGG	CTTATTTCCAA	AGCCAT	TCAGAGT	GGGATG	GCATAT	CCCTTAC	CTGGA	800	
Db	924	CAAGG	CTTATTTCCAA	AGCCAT	TCAGAGT	GGGATG	GCATAT	CCCTTAC	CTGGA	983	
Qy	801	GGCCCA	GATTTA	TGAGAA	GTGAGAC	CTGCGAGT	GTTCACAT	TTCTGT	GTGTAACA	860	
Db	984	GGCCCA	GATTTA	TGAGAA	GTGAGAC	CTGCGAGT	GTTCACAT	TTCTGT	GTGTAACA	1043	
Qy	861	TGCGT	CAGACTG	AGG	CCCTG	CTGAGT	GCCTGAGAGACAA	ACCCTCC	CAAGAGT	920	
Db	1044	TGCGT	CAGACTG	AGG	CCCTG	CTGAGT	GCCTGAGAGACAA	ACCCTCC	CAAGAGT	1103	
Qy	921	GACCT	CAAGC	CAAAAA	CAAA	AGCTTTCA	CTGAGT	GTGTA	TGCTTTCTTAA	980	
Db	1104	GACCT	CAAGC	CAAAAA	CAAA	AGCTTTCA	CTGAGT	GTGTA	TGCTTTCTTAA	1163	
Qy	981	TGAGC	CTTAGAT	CTATG	CTCAGAA	CACTTTAA	AGCAAT	CTCTC	ATATG	1040	
Db	1164	TGAGC	CTTAGAT	CTATG	CTCAGAA	CACTTTAA	AGCAAT	CTCTC	ATATG	1223	
Qy	1041	CATTA	ACCA	CAAGT	GTGGCTT	CTGCTG	CTCATG	AGAGAG	GCCTCTG	ATCTCAGTGG	1100
Db	1224	CATTA	ACCA	CAAGT	GTGGCTT	CTGCTG	CTCATG	AGAGAG	GCCTCTG	ATCTCAGTGG	1283

QY 1101 CTCGAACAGTCCCTTGCCCTCCATCTGATACAAAACATCTGACATCCCGCTCAGTA 1160
|
|
|
Db 1284 CTCGAACAGTCCCTTGCCCTCCATCTGATACAAAACATCTGACATCCCGCTCAGTA 1343
|
|
|
QY 1161 TTGGCACTTGTGGCTAATGAATACCTTCATGACAGCACTCCCTGACTGAATATCCGAGA 1220
|
|
|
Db 1344 TTGGCACTTGTGGCTAATGAATACCTTCATGACAGCACTCCCTGACTGAATATCCGAGA 1403
|
|
|
QY 1221 CAGTCTTGTGACCTTGTGAGATGTGTCTTGTGTGCTCCCTGACTGATACAGCTGTG 1280
|
|
|
Db 1404 CAGTCTTGTGACCTTGTGAGATGTGTCTTGTGTGCTCCCTGACTGATACAGCTGTG 1463
|
|
|
QY 1281 ATATCAGAGAGATGTGAGTACCTGTCTACTTCTATGAGTTTCGGACCCGCGCTCAGTG 1340
|
|
|
Db 1464 ATATCAGAGAGATGTGAGTACCTGTCTACTTCTATGAGTTTCGGACCCGCGCTCAGTG 1523
|
|
|
QY 1341 CTTTGAAGACAGAAACCCGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCGCTTTGT 1400
|
|
|
Db 1524 CTTTGAAGACAGAAACCCGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCGCTTTGT 1583
|
|
|
QY 1401 GTTCCGTGTGCTTCTCTGAAGGGGACATTTGTATGTTCGAAGAGCCAGCGAGAGA 1460
|
|
|
Db 1584 GTTCCGTGTGCTTCTCTGAAGGGGACATTTGTATGTTCGAAGAGCCAGCGAGAGA 1643
|
|
|
QY 1461 GAAAGTTACTGAGCCGGAAGATGATGAATACCTGGCTACCTTGTGCGAAGCGGAAATCC 1520
|
|
|
Db 1644 GAAAGTTACTGAGCCGGAAGATGATGAATACCTGGCTACCTTGTGCGAAGCGGAAATCC 1703
|
|
|
QY 1521 TAATGGGAACGACTGTCTCTGTGCGACAGCTTATATCTGACTGAGCACTACCTCAGCT 1580
|
|
|
Db 1704 TAATGGGAACGACTGTCTCTGTGCGACAGCTTATATCTGACTGAGCACTACCTCAGCT 1763
|
|
|
QY 1581 GGAATTGAACATGAGCTTGTGACAGAGACTCAAGAAACCGCGGTGAGATTTTGAACAG 1640
|
|
|
Db 1764 GGAATTGAACATGAGCTTGTGACAGAGACTCAAGAAACCGCGGTGAGATTTTGAACAG 1823
|
|
|
QY 1641 CACCATCCCGCTGAGCTGTCTGCGCCGACATGCTCCAGAGCTCTTCTTCTCTTAAC 1700
|
|
|
Db 1824 CACCATCCCGCTGAGCTGTCTGCGCCGACATGCTCCAGAGCTCTTCTTCTCTTAAC 1883
|
|
|
QY 1701 TTTCTCTCTCTCTCTCAGACCTTTCTTTTCTTTGTGTCTCTTGA 1746
|
|
|
Db 1884 TTTCTCTCTCTCTCTCAGACCTTTCTTTTCTTTGTGTCTCTTGA 1929
|
|
|

RESULT 8
US-10-104-047-249
; Sequence 249, Application US/10104047
; Publication No. US20030236392A1

; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR APPLICATION DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-249

Query Match 93.4%; Score 1631.2; DB 17; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;

Matches 1653; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 91 GGGCCTTCTGTGAAGGGCCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCAA 150
|
|
|
Db 138 GGGCCTTCTGTGAAGGGCCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCAA 197
|
|
|

QY 151 GTCACTGTGTGGGAAGCCCTGTGCTGAAAGTGTCTCTGAGTCCCTTTGTGCT 210
|
|
|
Db 198 GTCACTGTGTGGGAAGCCCTGTGCTGAAAGTGTCTCTGAGTCCCTTTGTGCT 257
|
|
|
QY 211 CCCCCGTGGGATCCCTGCGATTTACGAACCCGACGCTGATCGCCTTGGGATTAATTG 270
|
|
|
Db 258 CCCCCGTGGGATCCCTGCGATTTACGAACCCGACGCTGATCGCCTTGGGATTAATTG 317
|
|
|
QY 271 CGAAGAACCACTCTTACCTTAATT-----GTGCGTCCGAATCCAGAGTGGCT 320
|
|
|
Db 318 CGAAGAACCACTCTTACCTTAATTGTGAACAGAGTCCCTCCAGAACTCAGAGTGGCT 377
|
|
|
QY 321 GCTCTTAAGTCAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTG 380
|
|
|
Db 378 GCTCTTAAGTCAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTG 437
|
|
|
QY 381 CCTTACTGAAACATCTATGCGCTGCGCCAGCCGATACAGGCTCCAGCTCCCGTCTT 440
|
|
|
Db 438 CCTTACTGAAACATCTATGCGCTGCGCCAGCCGATACAGGCTCCAGCTCCCGTCTT 497
|
|
|
QY 441 GGTGTGTTCAGAGAGGTGCTTCAAGATGAGCTCAAGCTCCATCTTTGATGGGTCCG 500
|
|
|
Db 498 GGTGTGTTCAGAGAGGTGCTTCAAGATGAGCTCAAGCTCCATCTTTGATGGGTCCG 557
|
|
|
QY 501 CCTGGCTGCTATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
|
|
|
Db 558 CCTGGCTGCTATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
|
|
|
QY 561 TTTCTTCAACACATGAGGATCAGCATGCTCCGCGGAACTGCGGCTTCAAGAACCAAGTGC 620
|
|
|
Db 618 TTTCTTCAACACATGAGGATCAGCATGCTCCGCGGAACTGCGGCTTCAAGAACCAAGTGC 677
|
|
|
QY 621 TGTCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
|
|
|
Db 678 TGTCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
|
|
|
QY 681 CATCTTGGCGAGTCCGCGGAGCCATTAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 740
|
|
|
Db 738 CATCTTGGCGAGTCCGCGGAGCCATTAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 797
|
|
|
QY 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGCTGAGCATCATCTTACTCTGGA 800
|
|
|
Db 798 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGCTGAGCATCATCTTACTCTGGA 857
|
|
|
QY 801 GGGCCATGATTTATGAAGAGTGAAGACCTGAGGTGTGACATTTGTGTGTAAACA 860
|
|
|
Db 858 GGGCCATGATTTATGAAGAGTGAAGACCTGAGGTGTGACATTTGTGTGTAAACA 917
|
|
|
QY 861 TGGGTCAAGCTGTGAGGCTGCTGAGGTGCTGAGCAAAACCTCCAGAGGTGCT 920
|
|
|
Db 918 TGGGTCAAGCTGTGAGGCTGCTGAGGTGCTGAGCAAAACCTCCAGAGGTGCT 977
|
|
|
QY 921 GACCTCAGCCAGAAAAGTCTTTCATCTGAGTGTGATGAGTGTCTTCTTCTTA 980
|
|
|
Db 978 GACCTCAGCCAGAAAAGTCTTTCATCTGAGTGTGATGAGTGTCTTCTTCTTA 1037
|
|
|
QY 981 TGAAGCTCTAGATCTATGTGTCTCAGAAAAGCATTTAAAGCAATTCCTTCATCATGAGAT 1040
|
|
|
Db 1038 TGAAGCTCTAGATCTATGTGTCTCAGAAAAGCATTTAAAGCAATTCCTTCATCATGAGAT 1097
|
|
|
QY 1041 CAATTAACACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
|
|
|
Db 1098 CAATTAACACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
|
|
|
QY 1101 CTCGAACAGTCCCTTGCCCTCCATCTGATACAAAACATCTGACATCCCGCTCAGTA 1160
|
|
|
Db 1158 CTCGAACAGTCCCTTGCCCTCCATCTGATACAAAACATCTGACATCCCGCTCAGTA 1217
|
|
|
QY 1161 TTGGCACTTGTGGCTAATGAATACCTTCATGACAGCACTCCCTGACTGAATATCCGAGA 1220
|
|
|
Db 1218 TTGGCACTTGTGGCTAATGAATACCTTCATGACAGCACTCCCTGACTGAATATCCGAGA 1277
|
|
|

```

QY 1221 CAGCTTCTGAGACTGTTGGAGATGTTCTTTGGTCCCTGCACATGATCAAGCTCG 1280
Db 1278 CAGCTTCTGAGACTGTTGGAGATGTTCTTTGGTCCCTGCACATGATCAAGCTCG 1337
QY 1281 ATATCAAGAGATGCTGTGCACTGTCTACTATGATGTTTGGACCGGCTCAGTG 1340
Db 1338 ATATCAAGAGATGCTGTGCACTGTCTACTATGATGTTTGGACCGGCTCAGTG 1397
QY 1341 CTTTGAAGACAGAAAGCGGCTTTGTCAAGCCGACACGCTGATGAATGCTGCTTGT 1400
Db 1398 CTTTGAAGACAGAAAGCGGCTTTGTCAAGCCGACACGCTGATGAATGCTGCTTGT 1457
QY 1401 GTTCGTGTGCTCTTCTGAAAGGGGACATTTATGTTGAAAGAGACCGAGAGGA 1460
Db 1458 GTTCGTGTGCTCTTCTGAAAGGGGACATTTATGTTGAAAGAGACCGAGAGGA 1517
QY 1461 GAAGTTACTAGCCGGAAGATGATGAATACTGGGCTACCTTTTGCTGAAACCGGAATCC 1520
Db 1518 GAAGTTACTAGCCGGAAGATGATGAATACTGGGCTACCTTTTGCTGAAACCGGAATCC 1577
QY 1521 TAAATGGAAAGCACTGTCTGTGGCCAGCTTAATCTGACGAGACGATCTCCAGCT 1580
Db 1578 TAAATGGAAAGCACTGTCTGTGGCCAGCTTAATCTGACGAGACGATCTCCAGCT 1637
QY 1581 GGACTTGAACATGAGCTCGGACAGAGACTCAAAAGAACCGGGTGAATTTTGAACGAG 1640
Db 1638 GGACTTGAACATGAGCTCGGACAGAGACTCAAAAGAACCGGGTGAATTTTGAACGAG 1697
QY 1641 CACCATCCCCCTGATCTGTGCTGCCCTCCACATGCTCCCAAGTCTTTCTTCTTAAC 1700
Db 1698 CACCATCCCCCTGATCTGTGCTGCCCTCCACATGCTCCCAAGTCTTTCTTCTTAAC 1757
QY 1701 TTTCTCTCTCTCTCTCTCAACCTTTCTTTTCTTTTGTGCTCTTGA 1746
Db 1758 TTTCTCTCTCTCTCTCTCAACCTTTCTTTTCTTTTGTGCTCTTGA 1803

```

RESULT 9

```

US-10-114-270-195
/ Sequence 195, Application US/10114270
/ Publication No. US20040030110A1
/ GENERAL INFORMATION:
/ APPLICANT: Guo, Xiaojia
/ APPLICANT: Kehuda, Ramesh
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Malyankar, Uriel M.
/ APPLICANT: Spytek, Kimberly A.
/ APPLICANT: Patnurejan, Meera
/ APPLICANT: Liu, Ziaohong
/ APPLICANT: Guey, Vladimír Y.
/ APPLICANT: Li, Li
/ APPLICANT: Vermet, Corine
/ APPLICANT: Zetnussen, Bryan D.
/ APPLICANT: Gorman, Linda
/ APPLICANT: Shenoy, Suresh G.
/ APPLICANT: Pena, Carol E. A.
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Burgess, Catherine E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Shinkels, Richard A.
/ APPLICANT: Gangolli, Esna A.
/ APPLICANT: Taupier Jr., Raymond J.
/ APPLICANT: Caeman, Stacie J.
/ APPLICANT: Uí, Weizhen
/ APPLICANT: Anderson, David W.
/ APPLICANT: Lieté, Mario W.
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Edinger, Shlomo R.
/ APPLICANT: Stone, David J.
/ APPLICANT: Macdougall, John R.
/ APPLICANT: Rothenberg, Mark E.
/ TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same

```

```

/ FILE REFERENCE: 21402-322C
/ CURRENT APPLICATION NUMBER: US/10/114,270
/ CURRENT FILING DATE: 2002-11-27
/ PRIOR APPLICATION NUMBER: 60/281,086
/ PRIOR FILING DATE: 2001-04-03
/ PRIOR APPLICATION NUMBER: 60/281,136
/ PRIOR FILING DATE: 2001-04-03
/ PRIOR APPLICATION NUMBER: 60/281,863
/ PRIOR FILING DATE: 2001-04-05
/ PRIOR APPLICATION NUMBER: 60/281,906
/ PRIOR FILING DATE: 2001-04-05
/ PRIOR APPLICATION NUMBER: 60/282,020
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,930
/ PRIOR FILING DATE: 2001-04-10
/ PRIOR APPLICATION NUMBER: 60/282,934
/ PRIOR FILING DATE: 2001-04-10
/ PRIOR APPLICATION NUMBER: 60/283,512
/ PRIOR FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/283,710
/ PRIOR FILING DATE: 2001-04-13
/ PRIOR APPLICATION NUMBER: 60/284,234
/ PRIOR FILING DATE: 2001-04-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 470
/ SEQ ID NO 195
/ LENGTH: 1746
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1744)
US-10-114-270-195

```

Query Match 82.2%; Score 1435.6; DB 18; Length 1746;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1496; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

```

QY 81 GGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGAAACAGGCTGGATGATTTCA 140
Db 69 GGTCCAAAGAGGCGCTTCTGCTGAAGGCCACAGAGAAACAGGCTGGATGATTTCA 128
QY 141 GGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAAGCTTTCTTGGAGTCCC 200
Db 129 GGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAAGCTTTCTTGGAGTCCC 188
QY 201 CTTTGTCTCCCGCGCTGGATCCCTGCGATTTGAAGAACCGGACCTGCATGCGCTG 260
Db 189 CTTTGTCTCCCGCGCTGGATCCCTGCGATTTGAAGAACCGGACCTGCATGCGCTG 248
QY 261 GGATACTGGAGAAAGCACTCTTAACCTTAATTTGTGCTTCAGAACTCAGAGTGGCT 320
Db 249 GGATACTGGAGAAAGCACTCTTAACCTTAATTTGTGCTTCAGAACTCAGAGTGGCT 308
QY 321 GCTCTTAGATCAACATGCTCAAGGTGATTACCGAAATTTGGAGTGCAGAGACTG 380
Db 309 GCTCTTAGATCAACATGCTCAAGGTGATTACCGAAATTTGGAGTGCAGAGACTG 368
QY 381 CCTTAAGTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCAAGCTCCCGTCTT 440
Db 369 CCTTAAGTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCAAGCTCCCGTCTT 428
QY 441 GGTGTGTTCCAGAGAGGCTTCAAGCTGCTCAAGCTTCACTTTGATGGTCCGC 500
Db 429 GGTGTGTTCCAGAGAGGCTTCAAGCTGCTCAAGCTTCACTTTGATGGTCCGC 488
QY 501 CTTGCTGCTATAGAGAGCTGTGCTGTGTGCTGCTCAAGCTTCAAGCTTCAAGCTT 560
Db 489 CTTGCTGCTATAGAGAGCTGTGCTGTGTGCTGCTCAAGCTTCAAGCTTCAAGCTT 548
QY 561 TTTCTTCAACATGAGATGATGATGCTCCGGGAACTGGGCTTCAAGCAAGGTGCG 620
Db 549 TTTCTTCAACATGAGATGATGATGCTCCGGGAACTGGGCTTCAAGCAAGGTGCG 608

```


Db 493 TATGAGACGTGCTGTTGTGTCAGTACCGGCTAGGAATATTTGTTCTTACC 552
Qy 571 ACATGGATCAGCANTGCTCCGGGAACTGGGCTTTCAAGACAGGTGCTGCTTCC 630
Db 553 ACATGGATCAGCANTGCTCCGGGAACTGGGCTTTCAAGACAGGTGCTGCTTCC 612
Qy 631 TGGGTCCAGAAACATCGAGTTCTTGGTGGGACCAGCTGTCATCATTTTGGC 690
Db 613 TGGGTCCAGAAACATCGAGTTCTTGGTGGGACCAGCTGTCATCATTTTGGC 672
Qy 691 GAGTCCCGGGAGGCATTAAGTGTTCATGTTTACTGTCTCCATGSCCAAGCTTA 750
Db 673 GAGTCCCGGGAGGCATTAAGTGTTCATGTTTACTGTCTCCATGSCCAAGCTTA 732
Qy 751 TTCCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGGAGGCCCATGAT 810
Db 733 TTCCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGGAGGCCCATGAT 792
Qy 811 TATGAGAAAGTGAAGACCTGCAAGTGTTCACATTTCTGTGTAAACAATGCGTACAG 870
Db 793 TATGAGAAAGTGAAGACCTGCAAGTGTTCACATTTCTGTGTAAACAATGCGTACAG 852
Qy 871 TCTGAGGCCCCGTGAGGTGCTGAGAGACAAACCTCCCAAGAGCTGCTGACCTGACG 930
Db 853 TCTGAGGCCCCGTGAGGTGCTGAGAGACAAACCTCCCAAGAGCTGCTGACCTGACG 912
Qy 931 CAGAAAACAAAGTCTTCACTCGAGTGTGATGATGCTTCTTCTCTAATGAGCTCTA 990
Db 913 CAGAAAACAAAGTCTTCACTCGAGTGTGATGATGCTTCTTCTCTAATGAGCTCTA 972
Qy 991 GATCTATGCTCTAGAAAGCATTTAAAGCATTCCTTCATCATCGAGTCAATACAC 1050
Db 973 GATCTATGCTCTAGAAAGCATTTAAAGCATTCCTTCATCATCGAGTCAATACAC 1032
Qy 1051 GAGTGTGCTTCTGCTGCTAATGAAGAGGCTCCGAGATCTCAGTGGCTCCAAAG 1110
Db 1033 GAGTGTGCTTCTGCTGCTAATGAAGATTTGCTGCTCATACGCACTCCCTCA 1092
Qy 1111 TCCCTT-----GCCCTCATCTGATACAAACATCTCGACATC 1149
Db 1093 AACCGTATGACGCTTGGCTTCAACAGCTGGGCAATTCACAGAAACATCAGACATC 1152
Qy 1150 CCGCCTCAGTATTTGACCTTGCTGCTAATGAATCTTCATGACAGACATCCTGACT 1209
Db 1153 CCGCCTCAGTATTTGACCTTGCTGCTAATGAATCTTCATGACAGACATCCTGACT 1212
Qy 1210 GAAATCCGAGACAGTCTTCTGACCTTGGAGATGATGTTCTTGTGATCCCTGACATG 1269
Db 1213 GAAATCCGAGACAGTCTTCTGACCTTGGAGATGATGTTCTTGTGATCCCTGACATG 1272
Qy 1270 ATTCAGAGCTCATATCAGAGATCTGTGACCTGTCTAATCTTATGATTTCCGAC 1329
Db 1273 ATTCAGAGCTCATATCAGAGATCTGTGACCTGTCTAATCTTATGATTTCCGAC 1332
Qy 1330 CCGCCTCAGTCTTGAAGACAGAAAGCCGCTTTGTCAAGCCGACCAAGCTGATGAA 1389
Db 1333 CCGCCTCAGTCTTGAAGACAGAAAGCCGCTTTGTCAAGCCGACCAAGCTGATGAA 1392
Qy 1390 GTCCGCTTGTGTTGCTGCTGCTTCTGAAAGGGGACATTTGTTATGTTGAAGAGCC 1449
Db 1393 GTCCGCTTGTGTTGCTGCTGCTTCTGAAAGGGGACATTTGTTATGTTGAAGAGCC 1452
Qy 1450 ACGAGAGGAGAAAGTTACTGAGCCGAGAGATGATGAAATCTGCGCTACTTTGCTCGA 1509
Db 1453 ACGAGAGGAGAAAGTTACTGAGCCGAGAGATGATGAAATCTGCGCTACTTTGCTCGA 1512
Qy 1510 ACGGGAATCTTAATGGAAGCACTGTCTCTGTGCGACAGCTTAATCTGACAGAGCAG 1569
Db 1513 ACGGGAATCTTAATGGAAGCACTGTCTCTGTGCGACAGCTTAATCTGACAGAGCAG 1572
Qy 1570 TACTTCAAGCTGACATTGAACATGAGCTTGAGACAGAGACTCAAAAGACCGGAGTGGAT 1629
Db 1573 TACTTCAAGCTGACATTGAACATGAGCTTGAGACAGAGACTCAAAAGACCGGAGAGAT 1632

Qy 1630 TTTTGG 1635
Db 1633 GTGTGG 1638

RESULT 11
US-10-381-898-20
; Sequence 20, Application US/10381898
; Publication No. US2004008687A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valida; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKI, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; LAL, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANMALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
; APPLICANT: CHAWLA, Nandinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming; YUE, Henry
; APPLICANT: YANG, Junning; YAO, Montague; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7473645CB1
US-10-381-898-20

Query Match 79.4%; Score 1386.2; DB 18; Length 1857;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

Qy 81 GGGAAAAAAGTGGGCTTCTGCTGAAGGGCCACAGAGAAACACAGAGTGGATTC 140
Db 264 GGTCCAAAGAGGGCTTCTGCTGAAGGGCCACAGAGAAACACAGAGTGGATTC 323
Qy 141 GGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGGAGTCC 200
Db 324 GGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGGAGTCC 383
Qy 201 CTTTGCTGCTCCCGCTGGAGATCCCTGCGATTTACGAACCGGAGCCCTGATGCCCCCTG 260
Db 384 CTTTGCTGCTCCCGCTGGAGATCCCTGCGATTTACGAACCGGAGCCCTGATGCCCCCTG 443
Qy 261 GGATTAATTGCGAAGCACTCTTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCT 320

```
Db 444 GATTAATCTGGAGAGAGCACTCTCAACCTTAATTTGGCCCTCCAGAACTCAGAGTGGCT 503
Qy 321 GCTCTTAATCAACACATGCTCAAGGTGATTAACCCGAATTCGAGGTCAAGAGACTG 380
Db 504 GCTCTTAATCAACATATGCTCAAGGTGATTAACCCGAATTCGAGGTCAAGAGACTG 563
Qy 381 CCTCACTCAAGCAATCTATGAGGCTGCGCAAGCGATCAAGGCTCAAGGCTCCCGCTT 440
Db 564 CCTCACTCAAGCAATCTATGAGGCTGCGCAAGCGATCAAGGCTCAAGGCTCCCGCTT 623
Qy 441 GGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAAGCTTCAATCTTTGATGGTCCGC 500
Db 624 GGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAAGCTTCAATCTTTGATGGTCCGC 683
Qy 501 CTTGGCTGCTTATGAGACGTGCTGTGTGTGCTGTCAAGTACCGGCTGAGAAATTTGG 560
Db 684 CTTGGCTGCTTATGAGACGTGCTGTGTGTGCTGTCAAGTACCGGCTGAGAAATTTGG 743
Qy 561 TTTCTTCAACCATGAGATCAAGCATGCTCCGGGGAATCGGGCTTCAAGGACAGGTGGC 620
Db 744 TTTCTTCAACCATGAGATCAAGCATGCTCCGGGGAATCGGGCTTCAAGGACAGGTGGC 803
Qy 621 TGCTCTGTCTGGGTCCAGAAAGACATGAGTTCTTGGGTGGGACCCAGCTCTGTGAC 680
Db 804 TGCTCTGTCTGGGTCCAGAAAGACATGAGTTCTTGGGTGGGACCCAGCTCTGTGAC 863
Qy 681 CATCTTTGGCGAGTCCGGGGAGCCATTAAGTGTCTTATGCTTATATCTGTCTCCATGGC 740
Db 864 CATCTTTGGCGAGTCCGGGGAGCCATTAAGTGTCTTATGCTTATATCTGTCTCCATGGC 923
Qy 741 CAAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTGGGACATCCCTTACTGGA 800
Db 924 CAAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTGGGACATCCCTTACTGGA 983
Qy 801 GGCCCATGATTTATGAGAGAGTGAAGACCTGAGGTGTTGACATTTCTGTGTAAACA 860
Db 984 GGCCCATGATTTATGAGAGAGTGAAGACCTGAGGTGTTGACATTTCTGTGTAAACA 1043
Qy 861 TGCGTCAAGCTCTGAGGCTCTGCTGAGGTGCTGAGAACAAACCTTCCAGAGCTGCT 920
Db 1044 TGCGTCAAGCTCTGAGGCTCTGCTGAGGTGCTGAGAACAAACCTTCCAGAGCTGCT 1103
Qy 921 GACCTCGACCAAGAAACAAAGTCTTTGACTGAGTGTGATGAGTGTCTTCTTAA 980
Db 1104 GACCTCGACCAAGAAACAAAGTCTTTGACTGAGTGTGATGAGTGTCTTCTTAA 1163
Qy 981 TGAGGCTCTAGATCTATGCTCTCAAGAAAGCATTTAAAGCAATTCCTTCAATCGAGT 1040
Db 1164 TGAGGCTCTAGATCTATGCTCTCAAGAAAGCATTTAAAGCAATTCCTTCAATCGAGT 1223
Qy 1041 CAATTAACAGAGTGTGCTCTGCTGCTGCTTAAGAGAGGCTCTGAGATCTCTAGTG 1100
Db 1224 CAATTAACAGAGTGTGCTCTGCTGCTGCTTAAGAGAGGCTCTGAGATCTCTAGTG 1255
Qy 1101 CTCGCAAGAGTCCCTTGGCCCTTCAATCTGATCAAAACATCTCGACATCCCGCTCAGTA 1160
Db 1256 CTCGCAAGAGTCCCTTGGCCCTTCAATCTGATCAAAACATCTCGACATCCCGCTCAGTA 1274
Qy 1161 TTTGCACTTGGGTATGAAATATCTTCATGACAGACCTCCGACTGAATCCGAGA 1220
Db 1275 TTTGCACTTGGGTATGAAATATCTTCATGACAGACCTCCGACTGAATCCGAGA 1334
Qy 1221 CAGTCTTCTGACCTTGTGAGATGTGTTCTTTGTGTCTCTGACCTGATCAAGCTCG 1280
Db 1335 CAGTCTTCTGACCTTGTGAGATGTGTTCTTTGTGTCTCTGACCTGATCAAGCTCG 1394
Qy 1281 AATATCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340
Db 1395 AATATCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454
Qy 1341 CTTTGAAGACAGAGCGGCTTTTGTCAAGCCGACCAAGCTGATGAGTCCGCTTGT 1400
```

```
Db 1455 CTTTGAAGACAGAGCCAGCTTTTGTCAAAAGCCGACCAAGCTGATGAGTCCGCTTGT 1514
Qy 1401 GTTGGGTGTGCTTCTTCTGAAGGGGGACATTGTTATGTTGAGAGGACAGAGAGGA 1460
Db 1515 GTTGGGTGTGCTTCTTCTGAAGGGGGACATTGTTATGTTGAGAGGACAGAGAGGA 1574
Qy 1461 GAAGTTACTGAGCCCGAAGATGATGAAATTAATCTGAGCTACTTTGCTCAACCGGGAATCC 1520
Db 1575 GAAGTTACTGAGCCCGAAGATGATGAAATTAATCTGAGCTACTTTGCTCAACCGGGAATCC 1634
Qy 1521 TAATGGAAACCACTGTCTCTGTGGCCAGCTTATATCTGACTGACAGTACTTCAGCT 1580
Db 1635 TAATGGAAACCACTGTCTCTGTGGCCAGCTTATATCTGACTGACAGTACTTCAGCT 1694
Qy 1581 GACCTTGAACATGAGCCTTCGACAGAGACTCAAAAGACCGGGGGTGAATTTTGG 1635
Db 1695 GACCTTGAACATGAGCCTTCGACAGAGACTCAAAAGACCGGGAGAGATGTGTGG 1749

RESULT 12
US-10-094-749-736
; Sequence 736, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKI, ICHIRO
; APPLICANT: SEKI, NAOHKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOMYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 736
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-736

Query Match 77.0%; Score 1344.4; DB 17; Length 1962;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

Qy 91 GGGCTTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGCA 150
Db 222 GGGCTTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGCA 281
Qy 151 GTCACTGTGCTGGAAGGCTGTGCTGTGAAGGTCTCTCGAGTCCCTTTGCTGCT 210
Db 282 GTCACTGTGCTGGAAGGCTGTGCTGTGAAGGTCTCTCGAGTCCCTTTGCTGCT 341
Qy 211 CCCCCGCTGGAGATCCCTGCGATTTACGAACCCGACAGCTGATGCGCTCGGATTAATTG 270
Db 342 CCCCCGCTGGAGATCCCTGCGATTTACGAACCCGACAGCTGATGCGCTCGGATTAATTG 401
```


271 CGAAGACCACTCTACCTTAATTTGTCCTCAGAACTCAGAGTGTGCTTAAAT 330
 418 CGAAATGCCACATCTTACCTTAAATTATGCTTCAGAGCTTGAAGTGGTGTCTTAT 477
 331 CAACATCTCTCAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 390
 478 CAACAGTCTCTCAAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 537
 331 AACATCTATGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 450
 538 AACATCTATGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 597
 451 CCAGAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 510
 558 CCGGAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 657
 511 TATGAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 570
 658 TACGAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 717
 571 ACATGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 630
 718 ACAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 777
 631 TGGATCCAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 690
 778 TGGATCCAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 837
 691 GAGTCCGAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 750
 838 GAGTCCGAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 897
 751 TTCCCAAAAGCATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 807
 898 TTCCCAAAAGCATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 957
 808 GATTAATGAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 867
 958 GATTAATGAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1017
 868 GATTAATGAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 927
 1018 GATTAATGAGAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1077
 928 AGCCGAAAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 987
 1078 AGCCGAAAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1137
 988 CTAGATCTATGCTCAGAAAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1047
 1138 CTAGATCTATGCTCAGAAAGATCAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1197
 1048 CAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1107
 1138 CAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1254
 1108 AAGTCTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1167
 1255 AAGTCTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1314
 1168 CTTGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1227
 1315 CTTGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1374
 1228 CTTGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1287
 1375 CTTGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1434
 1288 AAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1347
 1435 AAGATGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1494
 1348 GACAGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1407

1495 GACAGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1554
 1408 GGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1467
 1555 GGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1614
 1468 CTTGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1527
 1615 CTTGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1674
 1528 AACAGCTGTCTGTGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1587
 1675 GAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1734
 1588 AACAGCTGTCTGTGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1646
 1735 AGTGTGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1794
 1647 -CCGCTGATCTGTGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1705
 1795 GTCCCTGATCTGTGAGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1854
 1706 TCTCTCTCTCGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1745
 1855 TTCTTGTCTCTCGAGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 1894

RESULT 14
 US-10-233-933A-1
 ; Sequence 1, Application US/10233933A
 ; Publication No. US2004021417A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamazaki, Teisuro
 ; APPLICANT: Miyazaki, Masao
 ; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
 ; FILE REFERENCE: SHIG F02U006
 ; CURRENT APPLICATION NUMBER: US/10/233, 933A
 ; PRIOR FILING DATE: 2002-09-03
 ; PRIOR APPLICATION NUMBER: JP2002-057908
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1629
 ; TYPE: DNA
 ; ORGANISM: Felis catus
 US-10-233-933A-1

Query Match 60.9%; Score 1063.2; DB 20; Length 1629;
 Best Local Similarity 80.9%; Pred. No. 0;
 Matches 165; Conservative 0; Mismatches 293; Indels 6; Gaps 2;
 91 GGGCTTCTGCTGAGAGGCAAGAGAAACAGAGTGTGAGATTCAGGGCAAGCA 150
 64 GGGCAAGCTGTGAGAGCAAGAGAGCAAGAGTGTGAGATTCAGGGCAAGCA 123
 151 GTACTGTGTGAGAGGCTGTGCTGTGAAGCTGTGCTGTGAGAGTGTGCTGT 210
 124 AACACTGTAGTGAAGCAAGCTGTGCTGTGAAGCTGTGCTGTGAGAGTGTG 183
 211 CCGGCTGTGAGTGTGCTGTGAGTGTGAGAAACAGAGTGTGAGATTCAGGGCAAGCA 270
 184 CTTCTGTAGAGGCTGTGAGTGTGAGAAACAGAGTGTGAGATTCAGGGCAAGCA 243
 271 CGAAGACCACTCTACCTTAATTTGTCCTCAGAACTCAGAGTGTGCTTAAAT 330
 244 CGAAATGCCACATCTTACCTTAAATTATGCTTCAGAGCTTGAAGTGGTGTCTTAT 303
 331 CAACATCTCTCAGAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 390
 304 CAACAGTCTCTCAAGTGTGCTTAACTCGAAATTCGAGTGTGAGAAAGCTGCTTAACTG 363

[illegible]

Db	1441	CTGAGCAGGAAGATGATGAGTACCTGGGCCAACTTTGCTCGGACC GGGA C C C T T A A C G G	1500
Oy	1528	AACGACTGTCTCTGTGGCCAGCTTATAATCTGACTGACAGTA C C T C A G C T T G A	1587
Db	1501	GAAAGTGTCCTCTGTGGCCAGCTCAACCCGAGACGAGCACTA C C T A A G C T G A T T T G	1566
Oy	1588	AACAATAGCCTGGAGAGAGACTCAAGAACC G C G G T G A T T T T G A C C A G C A C C A T C	1647
Db	1561	AGGTGATGGGTGGAGACAGAACTGAAGAGCAAGAGGTGGAGTTTGGATGAATACATT	1620
Oy	1648	CCCC 1651	
Db	1621	GTCC 1624	
 RESULT 15 US-10-451-168-42 Sequence 42, Application US/10451168 Publication No. US20040091969A1 GENERAL INFORMATION: APPLICANT: SMITHKLINE BEECHAM CORPORATION APPLICANT: GLAXO GROUP LIMITED TITLE OF INVENTION: NOVEL COMPOUNDS FILE REFERENCE: GP50039 CURRENT APPLICATION NUMBER: US/10/451,168 CURRENT FILING DATE: 2003-11-12 PRIOR APPLICATION NUMBER: PCT/US01/49232 PRIOR FILING DATE: 2000-12-17 PRIOR APPLICATION NUMBER: 60/256,710 PRIOR FILING DATE: 2000-12-19 PRIOR APPLICATION NUMBER: 60/257,048 PRIOR FILING DATE: 2000-12-20 PRIOR APPLICATION NUMBER: 60/260,482 PRIOR FILING DATE: 2001-01-09 PRIOR APPLICATION NUMBER: 60/264,922 PRIOR FILING DATE: 2001-01-30 PRIOR APPLICATION NUMBER: 60/266,797 PRIOR FILING DATE: 2001-02-06 PRIOR APPLICATION NUMBER: 60/276,988 PRIOR FILING DATE: 2001-03-19 PRIOR APPLICATION NUMBER: 60/281,535 PRIOR FILING DATE: 2001-04-04 PRIOR APPLICATION NUMBER: 60/289,622 PRIOR FILING DATE: 2002-06-28 NUMBER OF SEQ ID NOS: 110 SOFTWARE: FastSeq for Windows Version 4.0 SEQ ID NO 42 LENGTH: 1071 TYPE: DNA ORGANISM: Homo sapiens US-10-451-168-42			
 Query Match 58.5%; Score 1021.4; DB 18; Length 1071; Best Local Similarity 99.9%; Pred. No. 0; Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Oy	724	ATATCTGTCTCCCATGGCCAAAGGCTTATTTCCAAGGCATCATGAGAGTGGGGTGCC	783
Db	49	ATATCTGTCTCCCATGGCCAAAGGCTTATTTCCAAGGCATCATGAGAGTGGGGTGCC	108
Oy	784	ATATATCCCTTAACTGTGAGGSCCATGATTAATGAGAGATGAGAACCTGCAGCTGGTTGCA	843
Db	109	ATATATCCCTTAACTGTGAGGSCCATGATTAATGAGAGATGAGAACCTGCAGCTGGTTGCA	168
Oy	844	CATTTCTGTGTGAACAATGCGTCAGACTCTGAGGCCCTGTGAGGTGCTGAGGACAAAA	903
Db	169	CATTTCTGTGTGAACAATGCGTCAGACTCTGAGGCCCTGTGAGGTGCTGAGGACAAAA	228
Oy	904	CCCTCCAAAGACCTCTGACCTCTGACGACGAAAACMAAGCTTTTCACTCGAGTGGTTGAT	963
Db	229	CCCTCCAAAGACCTCTGACCTCTGACGACGAAAACMAAGCTTTTCACTCGAGTGGTTGAT	288


```
OY 964 GGTCCTTCTTCTTAATGAGCCTAGATCTATTGTCTCAGAAAGCATTTAAGCAATT 1023
    |||||
Db 289 GGTGCTTCTTCTTAATGAGCCTAGATCTATTGTCTCAGAAAGCATTTAAGCAATT 348
    |||||
OY 1024 CCTTCATCATCGAGTCATAATAACAAGTGTGCTTCTGCTGCTTAAGAAGAGCT 1083
    |||||
Db 349 CCTTCATCATCGAGTCATAATAACAAGTGTGCTTCTGCTGCTTAAGAAGAGAGCT 408
    |||||
OY 1084 CCTGAGATCTCTGAGTGTCTCCAAACAAGTCCCTTGCCTTCATCTGATACAAAACATCCCTG 1143
    |||||
Db 409 CCTGAGATCTCTGAGTGTCTCCAAACAAGTCCCTTGCCTTCATCTGATACAAAACATCCCTG 468
    |||||
OY 1144 CACATCCGCGCTCAGTATTGTGACCTTGTGGCTAAATGAATACTTCATGACAAGCACTCC 1203
    |||||
Db 469 CACATCCGCGCTCAGTATTGTGACCTTGTGGCTAAATGAATACTTCATGACAAGCACTCC 528
    |||||
OY 1204 CTGACTGAATCCGAGACAGTCTTCTGACCTTGTGAGATGTGTCTTGTGTGCTCCT 1263
    |||||
Db 529 CTGACTGAATCCGAGACAGTCTTCTGACCTTGTGAGATGTGTCTTGTGTGCTCCT 588
    |||||
OY 1264 GCACTGATCAAGCTCGATATCAAGAGATGCTGTGCACTGTCTACTTATGAGTTT 1323
    |||||
Db 589 GCACTGATCAAGCTCGATATCAAGAGATGCTGTGCACTGTCTACTTATGAGTTT 648
    |||||
OY 1324 CGGCAACCGGCGCTCAGTGTCTTGAAGACAGAAACCGGCTTTGTCAAGCCGACCAAGCT 1383
    |||||
Db 649 CGGCAACCGGCGCTCAGTGTCTTGAAGACAGAAACCGGCTTTGTCAAGCCGACCAAGCT 708
    |||||
OY 1384 GATGAAGTCCGCTTGTGTTCGGTGTGCTTCTGAAGGGGGAACATGTTAATGTTGAA 1443
    |||||
Db 709 GATGAAGTCCGCTTGTGTTCGGTGTGCTTCTGAAGGGGGAACATGTTAATGTTGAA 768
    |||||
OY 1444 GAGGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATAAATACTGGGCTACCTTT 1503
    |||||
Db 769 GAGGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATAAATACTGGGCTACCTTT 828
    |||||
OY 1504 GCTGGAACCGGGAATCTTAATGGGAAACGACCTGTCTGTGGCAAGCTTATATCTGACT 1563
    |||||
Db 829 GCTGGAACCGGGAATCTTAATGGGAAACGACCTGTCTGTGGCAAGCTTATATCTGACT 888
    |||||
OY 1564 GAGCAGTACCTTCAGCTGACTTGAACATGAGCTTCGACAGAGACTCAAGAACCGCGG 1623
    |||||
Db 889 GAGCAGTACCTTCAGCTGACTTGAACATGAGCTTCGACAGAGACTCAAGAACCGCGG 948
    |||||
OY 1624 GTGATTTTGGACAGACCAATCCCTGATCTGTCTGCTCCGACATGCTCCACAGT 1683
    |||||
Db 949 GTGATTTTGGACAGACCAATCCCTGATCTGTCTGCTCCGACATGCTCCACAGT 1008
    |||||
OY 1684 CCTCTTTCTTCTTAATCTTCTCTCTCCCTCCAGCCCTTTCTTTTCTTTGTGCTCCT 1743
    |||||
Db 1009 CCTCTTTCTTCTTAATCTTCTCTCTCCCTCCAGCCCTTTCTTTTCTTTGTGCTCCT 1068
    |||||
OY 1744 TGA 1746
    |||
Db 1069 TGA 1071
```

Search completed: June 15, 2005, 08:31:11
Job time : 1129.45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 7580.61 Seconds
(without alignment)
1160.423 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 1746
Sequence: 1 atgcacagggagctactc.....ttctcttgcctccttga 1746

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ha: *
2: gb_hg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_to: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1746	100.0	1746	6	AR438765 Sequence
2	1746	100.0	1746	6	AX600126 Sequence
3	1746	100.0	2158	6	AR438764 Sequence
4	1746	100.0	2158	6	AR600124 Sequence
5	1656	94.8	2047	9	BC069501 Homo sapi
6	1651.6	94.6	2232	6	AX477711 Sequence
7	1631.2	93.4	2092	9	AX746724 Sequence
8	1631.2	93.4	2092	9	AK090997 Homo sapi
9	1509.2	86.4	2298	9	BC039073 Homo sapi
10	1425	81.6	1716	6	CQ733044 Sequence
11	1386.2	79.4	1857	6	AX958392 Homo sapi
12	1344.4	77.0	1896	6	BC069548 Homo sapi
13	1344.4	77.0	1962	6	AX714052 Sequence
14	1344.4	77.0	1962	6	AK056109 Homo sapi
15	1198.4	66.6	1728	4	AB186392 Canis fam
16	1102	63.1	2139	4	AB045377 Felis cat
17	1041.2	59.6	1728	10	AB186393 Mus muscu
18	951.6	54.5	1739	9	AB060873 Macaca fa
19	634.6	36.3	921	6	AX380871 Sequence

20	413.8	23.7	2228	10	D50580
21	408.2	23.4	3391	9	AK095522
22	406.6	23.3	1846	9	D50579
23	406.6	23.3	2117	6	C0723952
24	406.6	23.3	2191	6	AX330319
25	406.6	23.3	2191	6	AX330320
26	406.6	23.3	2191	6	AX331596
27	406.6	23.3	2191	6	AX411210
28	406.6	23.3	2191	9	HSICR
29	406.6	23.2	3155	9	BC032095
30	405	23.2	1748	9	U60553
31	401.8	23.0	1831	10	BC028907
32	401.8	23.0	2139	9	BC053670
33	400.2	22.9	2114	9	AK000105
34	400.2	22.9	3824	6	AR528722
35	400.2	22.9	3824	6	AX358956
36	400.2	22.9	3824	6	AX362449
37	400.2	22.9	3824	6	AX454680
38	400.2	22.9	3824	6	AX464408
39	400.2	22.9	3824	6	AX491158
40	400.2	22.9	3824	9	AX358609
41	398.2	22.8	1641	6	AX359965
42	398.2	22.8	2087	6	AX359963
43	397.8	22.8	1848	10	D50577
44	397.6	22.8	2846	10	BC027185
45	396.2	22.7	2012	10	BC055062

ALIGNMENTS

RESULT 1
LOCUS AR438765 1746 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 3 from patent US 6664091.
ACCESSION AR438765
VERSION AR438765.1 GI:42663740
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1746)
AUTHORS Curtis,R.A.J. and Sillos-Santiago,I.
TITLE 53010, a human carboxylesterase family member and uses thereof
JOURNAL Patent: US 6664091-A 3 16-DEC-2003;
FEATURES
Location/Qualifiers
1..1746
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1746; DB 6; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCACAGGGAGCTTACTTCATCTGCTTCCAGATGATGCTTTCTGATTTCCAGCCC 60
DB 1 ATGCACAGGGAGCTTACTTCATCTGCTTCCAGATGATGCTTTCTGATTTCCAGCCC 60
QY 61 CTGTGGGACACAGACAGTGGGAAAACTGGGCTTTGCTGGAAGGACACAGAGAAC 120
DB 61 CTGTGGGACACAGACAGTGGGAAAACTGGGCTTTGCTGGAAGGACACAGAGAAC 120
QY 121 ACCAGGCTGGGATGATTCAGGACAGCAAGTCACTGTGCTGGGAAGCCCTGCTGTG 180
DB 121 ACCAGGCTGGGATGATTCAGGACAGCAAGTCACTGTGCTGGGAAGCCCTGCTGTG 180
QY 181 AACGTTCTTCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGATTTTCAAGAAC 240
DB 181 AACGTTCTTCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGATTTTCAAGAAC 240
QY 241 CCGCAGCTGCATCCCGCTGGGATTAATTGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300
DB 241 CCGCAGCTGCATCCCGCTGGGATTAATTGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 300

```

Db      241 CCGGAGCTGATGCGCCCTGGGATATACCTGGAGAGCAACCTCTAATTTGTGC 300
Qy      301 CTCGAGAACTCAGAGTGGCTGCTCTTAGATTAACAATGCTCAAGGTGATTAACCGAAA 360
Db      301 CTCGAGAACTCAGAGTGGCTGCTCTTAGATTAACAATGCTCAAGGTGATTAACCGAAA 360
Qy      361 TTGGAGATCAGAGAGCTGCTCTTACCTGAACATCTATGAGGCTGGCCAGCGCATACA 420
Db      361 TTGGAGATCAGAGAGCTGCTCTTACCTGAACATCTATGAGGCTGGCCAGCGCATACA 420
Qy      421 GGGCTCCAGATCCCGCTGTTGGTGGTTCACAGAGGTGCTTCAAGATGGCTGACGC 480
Db      421 GGGCTCCAGATCCCGCTGTTGGTGGTTCACAGAGGTGCTTCAAGATGGCTGACGC 480
Qy      481 TCCATCTTATGATGGATCCGCTGGCTGCTATGAGAGAGCTGCTGGTGTGTGCTCAG 540
Db      481 TCCATCTTATGATGGATCCGCTGGCTGCTATGAGAGAGCTGCTGGTGTGTGCTCAG 540
Qy      541 TACCGGCTAGAAATATTTGGTTTCTTACCAATGGGATCAGCATGCTCCGGGAACTGG 600
Db      541 TACCGGCTAGAAATATTTGGTTTCTTACCAATGGGATCAGCATGCTCCGGGAACTGG 600
Qy      601 GCGTTCAAGGACAGAGTGGCTGCTGCTGCTGGGTCAAGAGAAATCGAGTTCTTCGGT 660
Db      601 GCGTTCAAGGACAGAGTGGCTGCTGCTGCTGGGTCAAGAGAAATCGAGTTCTTCGGT 660
Qy      661 GGGGAGCCGAGCTGTGTGACCATCTTTGGCGAGTCCGCGGAGCCATAAGTTTCTAGT 720
Db      661 GGGGAGCCGAGCTGTGTGACCATCTTTGGCGAGTCCGCGGAGCCATAAGTTTCTAGT 720
Qy      721 CTTATATCTGTCTCCCATGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGTG 780
Db      721 CTTATATCTGTCTCCCATGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGTG 780
Qy      781 GGCATCATCTCTTACCTGAGAGGCCATGATTAAGAAGAGAGAGACCTGAGGTGGT 840
Db      781 GGCATCATCTCTTACCTGAGAGGCCATGATTAAGAAGAGAGAGACCTGAGGTGGT 840
Qy      841 GCACATTTCTGTGTAAACATGCGTCAACTCTGAGGCTGCTGAGGTGCTGAGAGACA 900
Db      841 GCACATTTCTGTGTAAACATGCGTCAACTCTGAGGCTGCTGAGGTGCTGAGAGACA 900
Qy      901 AAACCTCCAGAGAGCTGCTGAGCCCTCAGCCGAGAAAACAAAGTCTTTCACTCGAGTGT 960
Db      901 AAACCTCCAGAGAGCTGCTGAGCCCTCAGCCGAGAAAACAAAGTCTTTCACTCGAGTGT 960
Qy      961 GATGATGCTTTCTTTCTTAATGAGGCTCTAAGTCTATTGTCTCAGAAAGCATTTAAAGCA 1020
Db      961 GATGATGCTTTCTTTCTTAATGAGGCTCTAAGTCTATTGTCTCAGAAAGCATTTAAAGCA 1020
Qy      1021 ATTCTCTTCATCATGAGAGTCAATTAACAGAGATGTGGCTTCTGTGCTCATGAGAGAG 1080
Db      1021 ATTCTCTTCATCATGAGAGTCAATTAACAGAGATGTGGCTTCTGTGCTCATGAGAGAG 1080
Qy      1081 GCTCTGAGATCTCAAGTGGCTCCAAACAAAGTCCCTGGCCCTCATGATTAACAAACATC 1140
Db      1081 GCTCTGAGATCTCAAGTGGCTCCAAACAAAGTCCCTGGCCCTCATGATTAACAAACATC 1140
Qy      1141 CTGCAATATCCCGCTCAGATATTTGACCTTGTGCTAATGATATCTTCATGACAGAC 1200
Db      1141 CTGCAATATCCCGCTCAGATATTTGACCTTGTGCTAATGATATCTTCATGACAGAC 1200
Qy      1201 TCCCTGACTGAAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTCTTGTGTGTC 1260
Db      1201 TCCCTGACTGAAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTCTTGTGTGTC 1260
Qy      1261 CTTGCACTGATCAAGCTCGATATCAAGAGATGCTGTGCACTGTCTTCTTATGAG 1320
Db      1261 CTTGCACTGATCAAGCTCGATATCAAGAGATGCTGTGCACTGTCTTCTTATGAG 1320
Qy      1321 TTTTCGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTTGTCAAGCCGACAC 1380
Db      1321 TTTTCGGGACCGGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTTGTCAAGCCGACAC 1380

```

```

Qy      1381 GCTGATGAGTCCGCTTTGTGTTCGGTGTGCTTCTGAAGGGGAGACATTTATGTC 1440
Db      1381 GCTGATGAGTCCGCTTTGTGTTCGGTGTGCTTCTGAAGGGGAGACATTTATGTC 1440
Qy      1441 GAAGAGCCACGAGAGAGAGAAATTACTGAGCCGAAAGATGATAAATACTGGGCTAAC 1500
Db      1441 GAAGAGCCACGAGAGAGAGAAATTACTGAGCCGAAAGATGATAAATACTGGGCTAAC 1500
Qy      1501 TTTGCTCGAACCGGGATCTTAATGGGAAAGACCTGTCTGTGTGGCCGCTTAATATGTG 1560
Db      1501 TTTGCTCGAACCGGGATCTTAATGGGAAAGACCTGTCTGTGTGGCCGCTTAATATGTG 1560
Qy      1561 ACTGAGCAGTACTCTCAGCTGAGCTTGAACATGAGACCTCGAGACAGAGACTCAAGAACG 1620
Db      1561 ACTGAGCAGTACTCTCAGCTGAGCTTGAACATGAGACCTCGAGACAGAGACTCAAGAACG 1620
Qy      1621 GGGGTGAAATTTTGGACCAAGACCATCCCTGATCTGTGTGCTTCGACATGCTCAC 1680
Db      1621 GGGGTGAAATTTTGGACCAAGACCATCCCTGATCTGTGTGCTTCGACATGCTCAC 1680
Qy      1681 AGTCTCTTTCTCTTAACTTCTCTCTCTCTCCAGGCTTCTTTCTTTGTGCT 1740
Db      1681 AGTCTCTTTCTCTTAACTTCTCTCTCTCTCCAGGCTTCTTTCTTTGTGCT 1740
Qy      1741 CCTTGA 1746
Db      1741 CCTTGA 1746

RESULT 2
AX600126 1746 bp DNA linear PAT 14-FEB-2003
LOCUS AX600126
DEFINITION Sequence 3 from Patent WO250256.
ACCESSION AX600126
VERSION AX600126.1 GI:28400192
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Curtis, R.A. and Sillio-Santiago, I.
53010, a novel human carboxylesterase family member and uses
thereof
JOURNAL Patent: WO 0250256-A 3 27-JUN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1746; DB 6; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGCCACAGGAGCTTAATCTTCAATCTGCTTCAACATGATGTGCTTTTCTGATTTCCAGCC 60
Db      1 ATGCCACAGGAGCTTAATCTTCAATCTGCTTCAACATGATGTGCTTTTCTGATTTCCAGCC 60
Qy      61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
Db      61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
Qy      121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGGAAGCCCTGTGCTGTG 180
Db      121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGGAAGCCCTGTGCTGTG 180
Qy      181 AACGTTCTCTCGAGAGTCCCTTTGCTGCTCTCCCGCTGGAGATCCCTGCGATTTACGAC 240
Db      181 AACGTTCTCTCGAGAGTCCCTTTGCTGCTCTCCCGCTGGAGATCCCTGCGATTTACGAC 240

```


QY	241	CCGACGECCTGCATCGCCCTTGGGATAATCTTGCAGAAAGCCACTCTTACCTTAATTGTGC	300
Db	336	CCGACGACCCTGCATCGCCCTTGGGATAATCTTGCAGAAAGCCACTCTTACCTTAATTGTGC	395
QY	301	CTTCAGAACTCAGAGGCGCTGCTTTAGATCAACAATGCTCAAGTGCATTACCCGAAA	360
Db	396	CTTCAGAACTCAGAGTGGCTGCTTTAGATCAACAATGCTCAAGTGCATTACCCGAAA	455
QY	361	TTTCGAGTGTCAAGAACTGCTCTTACCTGAAACATCTATGCGCTGCACGCGCATACA	420
Db	456	TTTCGAGTGTCAAGAACTGCTCTTACCTGAAACATCTATGCGCTGCACGCGCATACA	515
QY	421	GGCTCAAGCTCCCGCTCTTGCTGTGTTTCCACAGAGGTGCTTTCAAGTGGCTCAGCC	480
Db	516	GGCTCAAGCTCCCGCTCTTGCTGTGTTTCCACAGAGGTGCTTTCAAGTGGCTCAGCC	575
QY	481	TCCATCTTTGATTTGGTCCGCTCGCTGCTATGAGACGTGCTGTTGTGTGCTCAG	540
Db	576	TCCATCTTTGATTTGGTCCGCTCGCTGCTATGAGACGTGCTGTTGTGTGCTCAG	635
QY	541	TACCGGCTTAGGAATTTTGTGTTTCTTACCAACAATGGGATCAGATGCTCCGGGAACTGG	600
Db	636	TACCGGCTTAGGAATAATTTGTGTTTCTTACCAACAATGGGATCAGATGCTCCGGGAACTGG	695
QY	601	GCTTTCAGAGACCAAGTGGCTGCTCTGTGCTGGGCTCCAGAAAGAAATCCGAGTTCTTCGT	660
Db	696	GCTTTCAGAGACCAAGTGGCTGCTCTGTGCTGGGCTCCAGAAAGAAATCCGAGTTCTTCGT	755
QY	661	GGGAGACCCCAAGCTCTGTGACCATTTTGTGGCGAGTCCGCGGGAGCCAAATGATGTTTCTAGT	720
Db	756	GGGAGACCCCAAGCTCTGTGACCATTTTGTGGCGAGTCCGCGGGAGCCAAATGATGTTTCTAGT	815
QY	721	CTTATACGTCTCCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG	780
Db	816	CTTATACGTCTCCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG	875
QY	781	GCCATCATCCCTTAACTTGAAGGCCCATGATTAATGAGAAAGTGAAGCACTGCAGGTGTT	840
Db	876	GCCATCATCCCTTAACTTGAAGGCCCATGATTAATGAGAAAGTGAAGCACTGCAGGTGTT	935
QY	841	GCACATTTCTGTGGTAAACAATTTGGTGTAGACTCTGAGAGCCCTGTGAGTGGCTTGAGACA	900
Db	936	GCACATTTCTGTGGTAAACAATTTGGTGTAGACTCTGAGAGCCCTGTGAGTGGCTTGAGACA	995
QY	901	AAACCTCCAGAGAGCTGTGACCTTCAGCCAGCAAGAAACAAAGTCTTCACTCGAGTGGT	960
Db	996	AAACCTCCAGAGAGCTGTGACCTTCAGCCAGCAAGAAACAAAGTCTTCACTCGAGTGGT	1055
QY	961	GATGGTGTCTTCTTCTCTAATGAGCCTTATGATCTTATGTCTCAAGAAAGCATTTAAACA	1020
Db	1056	GATGGTGTCTTCTTCTCTAATGAGCCTTATGATCTTATGTCTCAAGAAAGCATTTAAACA	1115
QY	1021	ATTCTTTCATCATTCGAGATGCAATAACAAGATGGCTTCTCTGTCGATGAAAGAG	1080
Db	1116	ATTCTTTCATCATTCGAGATGCAATAACAAGATGGCTTCTCTGTCGATGAAAGAG	1175
QY	1081	GGTCCCGAATCTCTAGTGGCTTCCACAAGTCCCTTGCCCTTCATCTGATACAAAACATC	1140
Db	1176	GGTCCCGAATCTCTAGTGGCTTCCACAAGTCCCTTGCCCTTCATCTGATACAAAACATC	1235
QY	1141	CTGCACATCCCGCTCAGTATTTGGAACCTGTGGCTATGAAATCTTCAAGACAAAGAC	1200
Db	1236	CTGCACATCCCGCTCAGTATTTGGAACCTGTGGCTATGAAATCTTCAAGACAAAGAC	1295
QY	1201	TCCCTGATCTGAAATCCGAGACAGTCTTTGGACTTTGCTTGGAGATGTGTTCTTTGTGTCT	1260
Db	1296	TCCCTGATCTGAAATCCGAGACAGTCTTTGGACTTTGCTTGGAGATGTGTTCTTTGTGTCT	1355
QY	1261	CCTGCATGATCAACGCTCGATATATCAAGAGATGTGGTGAACCTGTCTATCTTTATGAG	1320
Db	1356	CCTGCATGATCAACGCTCGATATATCAAGAGATGTGGTGAACCTGTCTATCTTTATGAG	1415
QY	1321	TTTTCGGACCGGCTCTCAGTGTCTTTGAAAGACAGAAAGCCGGCTTTTGTCTAAAGCCGATCAC	1380

Db 1416 TTTCGGCACCCGGCCTCAAGTCTTTGAAGACAGAAAGCCGGCTTTTGTCCAAAGCCGACAC 1475

Qy 1381 GCTGATGAAGTCCGCTTTGTGTGTCGGTGTGCTCTTCTGAAAGGGGACATTGTTATGTTTC 1440

Db 1476 GCTATGAAGTCCGCTTTGTGTGTCGGTGTGCTCTTCTGAAAGGGGACATTGTTATGTTTC 1535

Qy 1441 GAAGGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAATGATGAATACTCGGCTAAC 1500

Db 1536 GAAGGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAATGATGAATACTCGGCTAAC 1595

Qy 1501 TTTGCTCGAACCAGGGAATCTTAATGGAAGCACTGTCTCTGTGGCCAGCTTAATATCTG 1560

Db 1596 TTTCTCGAACCAGGGAATCTTAATGGAAGCACTGTCTCTGTGGCCAGCTTAATATCTG 1655

Qy 1561 ACTGAGCAGTACCTCCACAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1620

Db 1656 ACTGAGCAGTACCTCCACAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1715

Qy 1621 CGGATGATTTTGGACACAGACCATCCGCTGATCTGTCTGTCGCTCCGACATGCTCCAC 1680

Db 1716 CGGATGATTTTGGACACAGACCATCCGCTGATCTGTCTGTCGCTCCGACATGCTCCAC 1775

Qy 1681 AGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCAGGCTTTCTTTTCTTTTGTGCT 1740

Db 1776 AGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCAGGCTTTCTTTTCTTTTGTGCT 1835

Qy 1741 CCTTGA 1746

Db 1836 CCTTGA 1841

RESULT		4		
AX600124				
LOCUS	AX600124	2158 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0259256.			
ACCESSION	AX600124			PAT 14-FEB-2003
VERSION	AX600124.1	GI:28400190		
KEYWORDS	.	Homo sapiens (human)		
SOURCE		Homo sapiens		
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. .1		
REFERENCE	Curtis, R.A. and Silos-Santiago, I. 53010, a novel human carboxylesterase family member and uses thereof Patent: WO 0250256-A 1 27-JUN-2002;			
AUTHORS	Millemium Pharmaceuticals, Inc. (US) Location/Qualifiers			
TITLE	JOURNAL			
FEATURES				
source				
				1..2158

CDS

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
96..1841
/note="unamed protein product"
/codon_start=1
/db_xref="GI:28400191"
/translation="MPQGLTSSASQCEFLIQPLLGRQWQKTPSAGEPQNTRLG
MIGKQDQTVLGSPPVNPFLGVAPAPLPGSLRFTNPDPAPMNLRLASYPILCQ
NSELKLDQVHMKVHPKFGVSEBCLYINLYAPALDQSGSLPYLVHPGGAFTTGS
SITDGSFLAYEEDLVVVVQYRAGIFGFTTDDHAPENMAFKQVALSVQVQNIFF
FGDPSSTVYFGESAGIVSSVLIISPKAGIFHAIWESGVALIPLYEADYDEKSD
LQVAFHCNNAADSEALRICLRTPKSEELTISQKTSFPRVDGAFPPNRPDLIS
QKQFAPISIIQVNNHCEGFLIPMEKEPEILSGSKSLATHLIONIHPQYHIVA
NEBFHRSKLTVEIRDLDLGLGVFVVALTYRHDGAPFYFEPBPPPOCCD
TKFAPKADHAEVRFGGALKGDIVPFGATYRHDGAPFYFEPBPPPOCCD
GNLISMPAYNLTEOYLQDLNLMSSGQLKRPVDFWFTSTIPLILSSDMLHPSLSL
TFSLIQPFPPCAP"
```

Query Match

100.0%; Score 1746; DB 6; Length 2158;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCCACAGGGAATTAATCTATCTGCTTCACAATGAGTGTCTTTTCTGATTTCTCCAGCC 60
DB 96 ATGCCACAGGGAATTAATCTATCTGCTTCACAATGAGTGTCTTTTCTGATTTCTCCAGCC 155
QY 61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGGACAGAGAAC 120
DB 156 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGGACAGAGAAC 215
QY 121 ACCAGGCTGGGATGATTCAGGGGACAGAGTCACTGTGCTGGGAGGCCCTGTGCTGTG 180
DB 216 ACCAGGCTGGGATGATTCAGGGGACAGAGTCACTGTGCTGGGAGGCCCTGTGCTGTG 275
QY 181 AACGTTTCTCTGGAGTCCCTTTGCTGTCTCCCGCTGGGATCCCTGAGATTTACGAC 240
DB 276 AACGTTTCTCTGGAGTCCCTTTGCTGTCTCCCGCTGGGATCCCTGAGATTTACGAC 335
QY 241 CCGGAGCCGTCATGCGCCCTGGGATAACTTGGGAGGACCACTCCGTAACCTTAATTTGTC 300
DB 336 CCGGAGCCGTCATGCGCCCTGGGATAACTTGGGAGGACCACTCCGTAACCTTAATTTGTC 395
QY 301 CTCCAGAACTCAGAGTGGCTCTTTAGATCAACACATGCTGAGAGTGCATTAACCGAAA 360
DB 396 CTCCAGAACTCAGAGTGGCTCTTTAGATCAACACATGCTGAGAGTGCATTAACCGAAA 455
QY 361 TTCCGAGTGTCAAGAACTGCTCTTACCTGAACATCTATGCGCTGCCACCGCCATACA 420
DB 456 TTCCGAGTGTCAAGAACTGCTCTTACCTGAACATCTATGCGCTGCCACCGCCATACA 515
QY 421 GGGTCCAGTCCCGCTCTGTGTGTGTTCACAGAGGTCCTTAAGAATGGCTCAGCC 480
DB 516 GGGTCCAGTCCCGCTCTGTGTGTGTTCACAGAGGTCCTTAAGAATGGCTCAGCC 575
QY 481 TCCATCTTTGATGGTCCGCTGCTGCTGCTGATGAGGAGCTGCTGTTGTGCTCCAG 540
DB 576 TCCATCTTTGATGGTCCGCTGCTGCTGCTGATGAGGAGCTGCTGTTGTGCTCCAG 635
QY 541 TACCGGGTAGGAATTTGGTCTTTCACACACATGAGGATCAGCATGCTCCGGGAACTGG 600
DB 636 TACCGGGTAGGAATTTGGTCTTTCACACACATGAGGATCAGCATGCTCCGGGAACTGG 695
QY 601 GCGTTCAGAGACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 696 GCGTTCAGAGACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
QY 661 GGGGACCCGAGCTCTGTGACATCTTTGGCGAGTCCGCGGAGCCATTAAGTCTTCTAGT 720
DB 756 GGGGACCCGAGCTCTGTGACATCTTTGGCGAGTCCGCGGAGCCATTAAGTCTTCTAGT 815
QY 721 CTTATATCTGTCTCCATGCGCAAGAGCTTATTCACAAAGCCATCATGAGAGTGGGCTG 780
DB 816 CTTATATCTGTCTCCATGCGCAAGAGCTTATTCACAAAGCCATCATGAGAGTGGGCTG 875
QY 781 GGCATCATCTCTTACCTGAGGCGCCATGATTTATGAGAGAGTGAAGACTTGAAGTGTG 840
DB 876 GGCATCATCTCTTACCTGAGGCGCCATGATTTATGAGAGAGTGAAGACTTGAAGTGTG 935
QY 841 GCACATTTCTGTGTAAACAATGCTCAGACTCTGAGGCGCTGCTGAGGCTCTGAGAGCA 900
DB 936 GCACATTTCTGTGTAAACAATGCTCAGACTCTGAGGCGCTGCTGAGGCTCTGAGAGCA 995
QY 901 AAACCTTCAGAGAGCTGTGACCTCTCAGCGAGAAAACAAAGTCTTTCATCTGAGTGT 960
DB 996 AAACCTTCAGAGAGCTGTGACCTCTCAGCGAGAAAACAAAGTCTTTCATCTGAGTGT 1055
QY 961 GATGTGTCTTTTCTTCTTAATGAGCTCTGATCTAATTTGTCTCAGAGAACTTTTAAAG 1020
DB 1056 GATGTGTCTTTTCTTCTTAATGAGCTCTGATCTAATTTGTCTCAGAGAACTTTTAAAG 1115
QY 1021 ATTCTTTCATCATGAGAGTCAATAACAGAGTGTGCTTCTGCTGCTGCTTAAGAAGAG 1080

```

```

DB 1116 ATTCTTTCATCATGAGAGTCAATAACAGAGTGTGCTTCTGCTGCTGCTTAAGAAGAG 1175
QY 1081 GCTCTTGAGATCCCTGAGGCTCCCAAGATCCCTTGGCTCCATCTGATTAACAAATC 1140
DB 1176 GCTCTTGAGATCCCTGAGGCTCCCAAGATCCCTTGGCTCCATCTGATTAACAAATC 1235
QY 1141 CTGCACATCCCGCTCAGTATTTTGGACCTTGTGCTAATGATATCTTCATGACAGAC 1200
DB 1236 CTGCACATCCCGCTCAGTATTTTGGACCTTGTGCTAATGATATCTTCATGACAGAC 1295
QY 1201 TCCCTGACTGAATTCGAGACAGTCTTGTGACTTGTGAGATGTGTTCTTTGTGCTC 1260
DB 1296 TCCCTGACTGAATTCGAGACAGTCTTGTGACTTGTGAGATGTGTTCTTTGTGCTC 1355
QY 1261 CCGTGAATGATCAAGCTCGATATACAGAGTGTGTGACCTGTCTACTCTTATGAG 1320
DB 1356 CCGTGAATGATCAAGCTCGATATACAGAGTGTGTGACCTGTCTACTCTTATGAG 1415
QY 1321 TTTGGGACCCGCTCAGTGTCTTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1380
DB 1416 TTTGGGACCCGCTCAGTGTCTTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1475
QY 1391 GCTGATGAAGTCCGCTTGTGTGTGTGCTTCTTGAAGGGGACATTTATGTTT 1440
DB 1476 GCTGATGAAGTCCGCTTGTGTGTGTGCTTCTTGAAGGGGACATTTATGTTT 1535
QY 1441 GAGGAGCCACAGAGAGAGAGAGTACTGAGCCGGAAGATGATTAATCTGGGCTACC 1500
DB 1536 GAGGAGCCACAGAGAGAGAGAGTACTGAGCCGGAAGATGATTAATCTGGGCTACC 1595
QY 1501 TTTGCTGAAACCGGGAATCTTATGGAACGACTGTCTGTGTGACCATTAATCTG 1560
DB 1596 TTTGCTGAAACCGGGAATCTTATGGAACGACTGTCTGTGTGACCATTAATCTG 1655
QY 1561 ACTGAGCAGTACTCTCAGCTGACTTGAACATGAGCTCTGAGACAGACTCAAGAACCG 1620
DB 1656 ACTGAGCAGTACTCTCAGCTGACTTGAACATGAGCTCTGAGACAGACTCAAGAACCG 1715
QY 1621 CGGGTGAATTTTGGACCAAGACCATCCCTGATCTGTGCTCTCCGACATGCTCCAC 1680
DB 1716 CGGGTGAATTTTGGACCAAGACCATCCCTGATCTGTGCTCTCCGACATGCTCCAC 1775
QY 1681 AGTCTCTTTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 1776 AGTCTCTTTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
QY 1741 CCTTGA 1746
DB 1836 CCTTGA 1841

```

RESULT 5
BC069501
LOCUS
DEFINITION Homo sapiens hypothetical protein FLJ31547, mRNA (cDNA clone
MGC:97182 IMAGE:7262427), complete cds.
ACCESSION
BC069501
VERSION
BC069501.1 GI:46854682
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2047)
REFERENCE
AUTHORS
Klausner,R.D., Collins,F.S., Wagner,L., Shermen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buelow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedini,T.B., Toshitsugu,K.S.,
Carninci,P., Prange,C., Raha,S.S., Loguigliano,N.A., Peters,G.J.,
Abrahamson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, B., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J.,
 Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smailus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

2 (bases 1 to 2047)
 Director MGC Project.
 Direct Submission
 Submitted (29-APR-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Baylor Human Genome Sequencing Center
 cDNA Library Preparation: Baylor Human Genome Sequencing Center
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: am@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleaged, H.,
 Kowls, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRBR Plate: 4 Row: e Column: 1.
 Location/Qualifiers

1..2047
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:97182 IMAGE:7262427"
 /issue_type="PCR rescued clones"
 /clone_lib="NIH MGC 277"
 /note="Vector: pPCR-Script Amp SK(+) with reversed insert"
 1..2047
 /gene="PLJ31547"
 /db_xref="locusID:221223"
 121..1848
 /gene="PLJ31547"
 /codon_start=1
 /product="PLJ31547 protein"
 /protein_id="AAH69501.1"
 /db_xref="GI:46854683"
 /db_xref="locusID:221223"
 /translation="MSGNVHPGQILWAIWIAAPKGSARGPQNTLGIWQKO
 VTLGSPVNVFLGVPFAAPPLGSLFTNPQASPMNDLRATSYFNLCIQNSEWL
 LDQMLKRVHPKEGVSEDCVLYNTYAPAHADTGSKLPVLPFGAFAKTSASIPGS
 ALAAVEDLVVVVYRLGIFGFTTDOHAGNMAFDQVAAISWQKNIKEFGGPS
 SVITFGESAGAIIVSRLISLPMKGLPHKAIMESGVAITLYLRAHYEKSEDIQVAV
 FCGNNAADSEALICLARTKPSKELLTSCQKTPTRVDAFAFNEPLDLSQAKKA
 IPSIIGVNHCEGPIIPKRAPIRLSSNSLALHLONLHTIPQYLILVAAVEPHD
 KHSITETRDLDLDLGDVFVFPALITARYHRDGAAYVEYERFRHPQCEDEKPAFV
 KAHDAIDVRVFGAFLKGDIVMEGATBEESKLSRMKMYATFARTKPNQNDLSL
 WPAVNLTEQYLQDLNLSIGORLKEPRVDFWTSTIPLISASDMLSPISLTFSLIL
 QPFFFCAP"

ORIGIN

Query Match 94.8%; Score 1656; DB 9; Length 2047;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1656; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 GGCCCTTCTGCTGAAGGGCCAGAGGAACACAGGCTGGATGTGATTCAGGGCAAGAA 150
 Db 193 GGCCCTTCTGCTGAAGGGCCAGAGGAACACAGGCTGGATGTGATTCAGGGCAAGAA 252
 Qy 151 GTCACTGTCTGGGAAGCCCTGTGCTGAACTGTTCTCTGAGTCCCTTGTCTGCT 210
 Db 253 GTCACTGTCTGGGAAGCCCTGTGCTGAACTGTTCTCTGAGTCCCTTGTCTGCT 312
 Qy 211 CCCCCGCTGGGATCCCTGGAATTAACAAACCGAGCTGCATGCGCTGGGATTAATTG 270
 Db 313 CCCCCGCTGGGATCCCTGGAATTAACAAACCGAGCTGCATGCGCTGGGATTAATTG 372
 Qy 271 CGAAGAGCACCTCTTACCTTAATTTGAGCTCCAGAACTGAGTGGCTCTTATGAT 330
 Db 373 CGAAGAGCACCTCTTACCTTAATTTGAGCTCCAGAACTGAGTGGCTCTTATGAT 432
 Qy 331 CAACCATGCTCAAGTGCATTTACCCGAAATTTGAGTCAAGAACTGCTTACCTG 390
 Db 433 CAACCATGCTCAAGTGCATTTACCCGAAATTTGAGTCAAGAACTGCTTACCTG 492
 Qy 391 AACATCTATGCGCTGCGCCAGCGGATACAGGCTCCAAAGCTCCCGCTTGTGTGCTC 450
 Db 493 AACATCTATGCGCTGCGCCAGCGGATACAGGCTCCAAAGCTCCCGCTTGTGTGCTC 552
 Qy 451 CCAAGAGGCTGCTTCAAGACTGAGTCAAGCTCCATCTTTGATGGTCCGCTGCTGCGC 510
 Db 553 CCAAGAGGCTGCTTCAAGACTGAGTCAAGCTCCATCTTTGATGGTCCGCTGCTGCGC 612
 Qy 511 TATGAGACGCGTGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 570
 Db 613 TATGAGACGCGTGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
 Qy 571 ACATGGATCAGATGCTCCGAGGAACTGGGCTTCAAGAGCCAGAGTGGCTGCTGCTC 630
 Db 673 ACATGGATCAGATGCTCCGAGGAACTGGGCTTCAAGAGCCAGAGTGGCTGCTGCTC 732
 Qy 631 TGGGTCAGAAAGAAACATCGAGTCTTTCGGTGGGAGCCCGAGCTGTGACATCTTTGGC 690
 Db 733 TGGGTCAGAAAGAAACATCGAGTCTTTCGGTGGGAGCCCGAGCTGTGACATCTTTGGC 792
 Qy 691 GAGTCGCGGAGACCAATGAAGTGTTCAGTCTTATGCTGCTCCATGCGCAAGGCTTA 750
 Db 793 GAGTCGCGGAGACCAATGAAGTGTTCAGTCTTATGCTGCTCCATGCGCAAGGCTTA 852
 Qy 751 TTCCACAAAGCCATCATGGAAGTGGGTGGCCATCATCCTTACCTGAGGCCCATGAT 810
 Db 853 TTCCACAAAGCCATCATGGAAGTGGGTGGCCATCATCCTTACCTGAGGCCCATGAT 912
 Qy 811 TATGAGAGTGAAGGACCTGAGAGTGTGACATTTCTGTGTAAACAATGCTGACAGC 870
 Db 913 TATGAGAGTGAAGGACCTGAGAGTGTGACATTTCTGTGTAAACAATGCTGACAGC 972
 Qy 871 TCTGAGAGCTCTGAGTGTGCTGAGGACAAACCTCCAGAGAGCTGCTGCTGACG 930
 Db 973 TCTGAGAGCTCTGAGTGTGCTGAGGACAAACCTCCAGAGAGCTGCTGCTGACG 1032
 Qy 931 CAGAAACAAAGCTTTCATCTGAGTGTGATGCTGCTTCTTCTTCTTAATGAGCTCTTA 990
 Db 1033 CAGAAACAAAGCTTTCATCTGAGTGTGATGCTGCTTCTTCTTCTTAATGAGCTCTTA 1092
 Qy 991 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTCTTCATCATCGAGTCAATTAACAC 1050
 Db 1093 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTCTTCATCATCGAGTCAATTAACAC 1152
 Qy 1051 GAGTGTGCTTCTGCTGCTGCTGCTGAGAGGCTCTGAGATCTCAGTGGCTCCAAAG 1110
 Db 1153 GAGTGTGCTTCTGCTGCTGCTGCTGAGAGGCTCTGAGATCTCAGTGGCTCCAAAG 1212
 Qy 1111 TCCCTTGGCTTCATCTGATCAAAACATCTCTGCAATCCCGCTGATATTGACACTT 1170
 Db 1213 TCCCTTGGCTTCATCTGATCAAAACATCTCTGCAATCCCGCTGATATTGACACTT 1272

1171 GTGGCTAATGAACTATCTTCATGACAGCACTCCCTGACTGAATCCGAGACAGCTCTTCG 1230
1273 GTGGCTAATGAACTATCTTCATGACAGCACTCCCTGACTGAATCCGAGACAGCTCTTCG 1332
1231 GACTTGTCTTGAAGATGATGTTCTTTGTGTGCTCCCTGCACTGATCAAGCTGATATCAAGA 1290
1333 GACTTGTCTTGAAGATGATGTTCTTTGTGTGCTCCCTGCACTGATCAAGCTGATATCAAGA 1392
1231 GATGCTGTGATGACCTCTGTCTACTCTTATAGATTTCCGACCCGGCTCAAGCTCTTGAAGAC 1350
1393 GATGCTGTGATGACCTCTGTCTACTCTTATAGATTTCCGACCCGGCTCAAGCTCTTGAAGAC 1452
1351 AGCAAGCCGGCTTTTGTCAAGCCGACCACTGATGAAAGTCCGCTTTGTGTGCTGAT 1410
1453 AGCAAGCCGGCTTTTGTCAAGCCGACCACTGATGAAAGTCCGCTTTGTGTGCTGAT 1512
1411 GCTTCTCTGAAGGGGACATTTGTTATGTTTGAAGAGCCGACGAGAGAGAGAGATTACTG 1470
1513 GCTTCTCTGAAGGGGACATTTGTTATGTTTGAAGAGCCGACGAGAGAGAGATTACTG 1572
1471 AGCCGGAAGATGATGAAATATCTGGGCTACTTTGCTCCGACCCGGAAATCTTAATGGGAA 1530
1573 AGCCGGAAGATGATGAAATATCTGGGCTACTTTGCTCCGACCCGGAAATCTTAATGGGAA 1632
1531 GACTGTCTGTGTGGCCAGCTTATATCTGACCTGACAGTACTCTGAGCTGGAATTTGAAC 1590
1633 GACTGTCTGTGTGGCCAGCTTATATCTGACCTGACAGTACTCTGAGCTGGAATTTGAAC 1692
1591 ATGAGCTCTGGAACAGAGACTCAAGAAACCGCGGGTGGATTTTGAACAGCAACATCCGC 1650
1693 ATGAGCTCTGGAACAGAGACTCAAGAAACCGCGGGTGGATTTTGAACAGCAACATCCGC 1752
1651 CTGATCTCTGTGTGGCTCCGACATGCTCCAGATCTCTTTCTTCTTAATCTTCTCTCT 1710
1753 CTGATCTCTGTGTGGCTCCGACATGCTCCAGATCTCTTTCTTCTTAATCTTCTCTCT 1812
1711 CTCTCCAGAGCTTTCTTTTCTTTTGTGTGCTCTTGA 1746
1813 CTCTCCAGAGCTTTCTTTTCTTTTGTGTGCTCTTGA 1848

RESULT 6
AX477711 2232 bp DNA linear part 12-AUG-2002
DEFINITION Sequence 23 from Patent W00246426.
ACCESSION AX477711 GI:22216866
VERSION AX477711.1 GI:22216866
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 Sanjanwala, M.M., Yao, M.G., Au-Young, J., Baughn, M.R., Arvizu, C.,
Ring, H.Z., Lee, B.A., Ding, L., Hafalia, A.J., Tang, Y.T., Yue, H.,
Tribouley, C.M., Lu, D.A., Lai, P.G., Warren, B.A., Yang, J.,
Wallis, N.K., Nguyen, D.B., Gandhi, A.R. and Ison, C.H.
TITLE Drug metabolizing enzymes
Patent: WO 0246426-A 23 13-JUN-2002;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
Source 1.2232
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 6538080CB1"

Query Match 94.6%; Score 1651.6; DB 6; Length 2232;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 81 GGGAAAACTGGGCTTCTGTGAAGGGCCACAGAGAAACACAGGCTGGATGATTC 140

264 GGTCCAAAGAGGCTTCTGTGAAGGGCCACAGAGAAACACAGGCTGGATGATTC 323
141 GGGCAAGCAATCACTGTGTCTGGAGAGCCCTGTGCTTGAACGTTCTCCGAGTCCC 200
324 GGGCAAGCAATCACTGTGTCTGGAGAGCCCTGTGCTTGAACGTTCTCCGAGTCCC 383
201 CTTGTCTGTCTCCCGCTGGAGATCCCTGCAATTTACGAACCCGACGCTGCATCCGCTG 260
384 CTTGTCTGTCTCCCGCTGGAGATCCCTGCAATTTACGAACCCGACGCTGCATCCGCTG 443
261 GGTAACTTGGAGAGAGCCACTCTTCACTTAATTTGTCTCCAGAACTCAGAGTGGCT 320
444 GGATTAATTTGGAGAGAGCCACTCTTCACTTAATTTGTCTCCAGAACTCAGAGTGGCT 503
321 GCTCTTAATCAACATGCTCAAGTGCATTAACCCGAAATTCGAGAGTCAAGAGACTG 380
504 GCTCTTAATCAACATGCTCAAGTGCATTAACCCGAAATTCGAGAGTCAAGAGACTG 563
381 CCTTACCTGAAACATCTATGAGGCTTCCGACGAGATACAGGCTCCAGGCTCCGCTT 440
564 CCTTACCTGAAACATCTATGAGGCTTCCGACGAGATACAGGCTCCAGGCTCCGCTT 623
441 GGTGTGTTCCTCCAGAGAGTGCCTTCAAGACTGCTCAAGCTTCAATCTTTGATGGTCCG 500
624 GGTGTGTTCCTCCAGAGAGTGCCTTCAAGACTGCTCAAGCTTCAATCTTTGATGGTCCG 683
501 CTTGGCTGCTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
684 CTTGGCTGCTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743
561 TTTCTTCAACATGAGATTCAGATGCTCCGAGGAACTGGGCTTCAAGAACAGAGTGGC 620
744 TTTCTTCAACATGAGATTCAGATGCTCCGAGGAACTGGGCTTCAAGAACAGAGTGGC 803
621 TGTCTGTCTGTGGGTCAAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTAC 680
804 TGTCTGTCTGTGGGTCAAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTAC 863
681 CATCTTTGGAGAGTCCGCGGAGACATAGTGTCTTATGCTTATGCTCTCCATAGGC 740
864 CATCTTTGGAGAGTCCGCGGAGACATAGTGTCTTATGCTTATGCTCTCCATAGGC 923
741 CAAGGCTTATTCACAAAGCATATGAGAGTGGGAGCATATCCCTTACCTGGA 800
924 CAAGGCTTATTCACAAAGCATATGAGAGTGGGAGCATATCCCTTACCTGGA 983
801 GGCCCATGATTAAGAGAGTGAAGACTGAGAGTGTGTCATTTCTGTGTAAACA 860
984 GGCCCATGATTAAGAGAGTGAAGACTGAGAGTGTGTCATTTCTGTGTAAACA 1043
861 TGGCTCAAGCTTGTAGAGCCCTGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 920
1044 TGGCTCAAGCTTGTAGAGCCCTGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1103
921 GACCTTCAAGCCAGAAACAAAGTCTTTGATCTGAGTGTGTGTGTGTGTGTGTGTGTGT 980
1104 GACCTTCAAGCCAGAAACAAAGTCTTTGATCTGAGTGTGTGTGTGTGTGTGTGTGTGT 1163
981 TGAAGCTTCAAGTCTATTTGTCTCAGAAAGCAATTTAAACAAATCTTCCATCATCGAGT 1040
1164 TGAAGCTTCAAGTCTATTTGTCTCAGAAAGCAATTTAAACAAATCTTCCATCATCGAGT 1223
1041 CAATTAACAAGAGT 1100
1224 CAATTAACAAGAGT 1283
1101 CTCCAAAGTCCCTTGGCTTCCATCTGATTAACAAACATCTCCGACATCCGCTCAGTA 1160
1284 CTCCAAAGTCCCTTGGCTTCCATCTGATTAACAAACATCTCCGACATCCGCTCAGTA 1343
1161 TTTGCACTTGT 1220

Db 1344 TTGCACTTGTGGCTAATGTAATCTTCATGACAAGCACTCCGTGACTGAATCCGGA 1403
Qy 1221 CAGTCTTCTGGACTTGCTGGAGATGTTCTTTGTGGTCCCTGCACTGATCAGCTCG 1280
Db 1404 CAGTCTTCTGGACTTGCTGGAGATGTTCTTTGTGGTCCCTGCACTGATCAGCTCG 1463
Qy 1281 AATATCAGAGATGCTGTGTCACCTGTCTACTTATGATGTTTGGCACTGGCCCTGAGTG 1340
Db 1464 AATATCAGAGATGCTGTGTCACCTGTCTACTTATGATGTTTGGCACTGGCCCTGAGTG 1523
Qy 1341 CTTTGAAGCAGAGAGCCGGCTTTTGTGAAGCCGACACGCTGATGAAGTCCGCTTGT 1400
Db 1524 CTTTGAAGCAGAGAGCCGGCTTTTGTGAAGCCGACACGCTGATGAAGTCCGCTTGT 1583
Qy 1401 GTTGGTGTGCTCTCTGGAAGGGGACATTTGTTATGTTGAAGAGCCAGAGAGA 1460
Db 1584 GTTGGTGTGCTCTCTGGAAGGGGACATTTGTTATGTTGAAGAGCCAGAGAGA 1643
Qy 1461 GAAGTTACTGAGCCGGGAAGATGTAATAATCTGGGCTTACCTTTGCTGAAACGGGAATCC 1520
Db 1644 GAAGTTACTGAGCCGGGAAGATGTAATAATCTGGGCTTACCTTTGCTGAAACGGGAATCC 1703
Qy 1521 TAATGGGAACGACCTGTCTGTGGCCAGCTTAATCTGACGAGAGTACTCCAGCT 1580
Db 1704 TAATGGGAACGACCTGTCTGTGGCCAGCTTAATCTGACGAGAGTACTCCAGCT 1763
Qy 1581 GGAATTGAACATGAGCTCTCGACAGAGACTCAAGAACCGGGGTGGAATTTTGAACAG 1640
Db 1764 GGAATTGAACATGAGCTCTCGACAGAGACTCAAGAACCGGGGTGGAATTTTGAACAG 1823
Qy 1641 CACCATCCCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTTTTCTTAAAC 1700
Db 1824 CACCATCCCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTTTTCTTAAAC 1883
Qy 1701 TTTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCTCTGA 1746
Db 1884 TTTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCTCTGA 1929

RESULT 7
AX746724 2092 bp mRNA linear PAT 20-JUN-2003
LOCUS AX746724
DEFINITION Sequence 249 from Patent EPI308459.
ACCESSION AX746724
VERSION AX746724.1 GI:32130991
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Iwoga, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
Maehuo, Y.
TITLE Full-length cDNA sequences
Patent: EP 1308459-A 249 07-MAY-2003;
JOURNAL Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
FEATURES
Source location/Qualifiers
1.2092
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 93.4%; Score 1631.2; DB 6; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 3; Indels 10; Gaps 1;
Qy 91 GGGCTTCTCTGTAAGGGCCACAGAGAACACCAAGGCTGGATGATTCAGGGCAGAGCA 150
Db 138 GGGCTTCTCTGTAAGGGCCACAGAGAACACCAAGGCTGGATGATTCAGGGCAGAGCA 197

Qy 151 GTGACGTGTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGAGAGTCCCTTGTGCT 210
Db 198 GTGACGTGTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGAGAGTCCCTTGTGCTGCT 257
Qy 211 CCCCCTGGGATTCCTGCGATTTACGAACCCGACGCTGACATGCGCCCTGGGATTAAGT 270
Db 258 CCCCCTGGGATTCCTGCGATTTACGAACCCGACGCTGACATGCGCCCTGGGATTAAGT 317
Qy 271 CGAAGAGCACTCTTACCTTAATTT-----GTGCTTCAGAACTCAGAGTGGCT 320
Db 318 CGAAGAGCACTCTTACCTTAATTTGTAAGACAGAGTGGCTTCAGAACTCAGAGTGGCT 377
Qy 321 GCTTTAGATCAACATGCTCAAGTGCATTAACCCGAATTCGAGAGTGTACGAAGCTG 380
Db 378 GCTTTAGATCAACATGCTCAAGTGCATTAACCCGAATTCGAGAGTGTACGAAGCTG 437
Qy 381 CCTTACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTCAAGCTCCCGCTT 440
Db 438 CCTTACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTCAAGCTCCCGCTT 497
Qy 441 GGTGTGTTCCAGAGAGGTGCTTCAAGATGCTGACCTTCCATCTTTGATGAGTCCG 500
Db 498 GGTGTGTTCCAGAGAGGTGCTTCAAGATGCTGACCTTCCATCTTTGATGAGTCCG 557
Qy 501 CTTGGCTGCTATAGGAAGTGTGTTGTGTCGTCAGTACCGGCTAGAAATTTTGG 560
Db 558 CTTGGCTGCTATAGGAAGTGTGTTGTGTCGTCAGTACCGGCTAGAAATTTTGG 617
Qy 561 TTTCTTACCAACATGGGATGAGATGCTCCGGGAACTGGGCTTCAAGACAGGTGCG 620
Db 618 TTTCTTACCAACATGGGATGAGATGCTCCGGGAACTGGGCTTCAAGACAGGTGCG 677
Qy 621 TGTCTGTCTGTGGTCCAGAAACATCGAGTTCTTGGGAGGACCCAGGCTGTGAC 680
Db 678 TGTCTGTCTGTGGTCCAGAAACATCGAGTTCTTGGGAGGACCCAGGCTGTGAC 737
Qy 681 CATCTTTGGGAGTCCCGGGAGCCATTAAGTGTCTTATATGTCCTCCATGGC 740
Db 738 CATCTTTGGGAGTCCCGGGAGCCATTAAGTGTCTTATATGTCCTCCATGGC 797
Qy 741 CAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGGCAATCCCTTACTTGA 800
Db 798 CAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGGCAATCCCTTACTTGA 857
Qy 801 GGCCCATGATTAAGAGAGAGAGAGAGCTGCAAGTGTGACATTTCTGTGTAACA 860
Db 858 GGCCCATGATTAAGAGAGAGAGAGAGCTGCAAGTGTGACATTTCTGTGTAACA 917
Qy 861 TGGTTCAGACTGAGAGCCCTGCTGAGAGTGCCTGAGCAAAACCTTCAGAGAGTGGCT 920
Db 918 TGGTTCAGACTGAGAGCCCTGCTGAGAGTGCCTGAGCAAAACCTTCAGAGAGTGGCT 977
Qy 921 GACCTTCAGCCGAAAGCAAAAGTCTTTTCACTGAGTGTGATGATGCTTTCTTCTAA 980
Db 978 GACCTTCAGCCGAAAGCAAAAGTCTTTTCACTGAGTGTGATGATGCTTTCTTCTAA 1037
Qy 981 TGAGCCTTAAGATCAATGCTCAGAAAGCATTTTAAGCAATTCCTTCATCGAGAGT 1040
Db 1038 TGAGCCTTAAGATCAATGCTCAGAAAGCATTTTAAGCAATTCCTTCATCGAGAGT 1097
Qy 1041 CAATTAACAGAGTGTGCTTCTGTGCTGCTATGAGAGAGCTCTGAGATCCCTGAGTGG 1100
Db 1098 CAATTAACAGAGTGTGCTTCTGTGCTGCTATGAGAGAGCTCTGAGATCCCTGAGTGG 1157
Qy 1101 CTCGAAGAGTCCCTTGGCCCTTCATCTGATTAACAAACATCTTGCATATCCGCTGAGTA 1160
Db 1158 CTCGAAGAGTCCCTTGGCCCTTCATCTGATTAACAAACATCTTGCATATCCGCTGAGTA 1217
Qy 1161 TTTGCACTTGTGCTATGATTAATCTTCCATGAACAAGCACTCCCTGATGAAATCCGAGA 1220
Db 1218 TTTGCACTTGTGCTATGATTAATCTTCCATGAACAAGCACTCCCTGATGAAATCCGAGA 1277

Qy	1221	TAAGCTTCCTGGA	CTTTCGCTTGGAG	AGTGTCTCTTTGTGGT	CCCTGAC	CTGATCA	CAGAGCTG	1280	
Db	1278	CAGCTTCCTTGA	CTTGTGGAG	AGTGTCTTTGTGGT	CCCTCACT	CACTGATCA	CAGAGCTCG	1337	
Qy	1281	ATTATCACAG	AATGCTCGGTG	CAACCTGTCTTA	CTTCTATG	ATGATTTTG	GGCACCGGCTCTAGTG	1340	
Db	1338	ATTATCACAG	AGATGCTGGT	GTGCACTGTCTACTT	CTATG	AGTTTGG	CACCGGCTCTAGTG	1397	
Qy	1341	CTTTGAAG	ACAGAA	CGCGGCTTTTGTCA	AAAGCCG	ACACGCTG	ATGAAAGTCGCTTTGT	1400	
Db	1398	CTTTGAAG	ACAGAA	CGCGGCTTTTGTCA	AAAGCCG	ACACGCTG	ATGAAAGTCGCTTTGT	1457	
Qy	1401	GTTTGGTGTGT	GCTCTTCTCTGA	AAAGGGG	AGACATTTG	TATGTTG	TGCAAGAAGCTCA	CGAGAGAGA	1460
Db	1458	GTTTGGTGTGT	GCTCTTCTCTGA	AAAGGGG	AGACATTTG	TATGTTG	TGCAAGAAGCTCA	CGAGAGAGA	1517
Qy	1461	GAGCTTACTG	AGCCGG	GAAGATG	ATGAATATCTGG	GGCTACCTTTG	CTCGAA	CCGGGAATCC	1520
Db	1518	GAGCTTACTG	AGCCGG	GAAGATG	ATGAATATCTGG	GGCTACCTTTG	CTCGAA	CCGGGAATCC	1577
Qy	1521	TAATGGG	ACAACACCTGTCTCTGTGTGG	CCAGCTTATATCTGA	CTAGAGCA	GTACCTCA	GCT	1580	
Db	1578	TAATGGG	ACAACACCTGTCTCTGTGTGG	CCAGCTTATATCTGA	CTAGAGCA	GTACCTCA	GCT	1637	
Qy	1581	GGACTTGA	ACAATGAGCTCTCG	ACAGAGA	CTCAAA	GAACCGCGGGT	TGATTTT	TGACACAG	1640
Db	1638	GGACTTGA	ACAATGAGCTCTCG	ACAGAGA	CTCAAA	GAACCGCGGGT	TGATTTT	TGACACAG	1697
Qy	1641	CACCATCC	CCCTGATCTGTGTGCTCTCG	ACATGCTCA	CAGTCTCTTTTCTT	CCTTAAC		1700	
Db	1698	CACCATCC	CCCTGATCTGTGTGCTCTCG	ACATGCTCA	CAGTCTCTTTTCTT	CCTTAAC		1757	
Qy	1701	TTTCTCTCTCT	CTCCAGAGCTTTCTTTTCTTTT	TGTGCTCTCTGA				1746	
Db	1758	TTTCTCTCTCT	CTCCAGAGCTTTCTTTTCTTTT	TGTGCTCTCTGA				1803	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS
AKO90997	AKO90997	2092 bp	mRNA	linear	PRI 30-JAN-2004		
AKO90997	Homo sapiens cDNA FLJ31678 f1s, clone BAWH2002191, weakly similar to PAPPY ACY1-COA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).	AKO90997					
AKO90997	AKO90997.1 GI:21749268						
AKO90997	oligo capping: f1s (full insert sequence).						
AKO90997	Homo sapiens (human)						
AKO90997	Homo sapiens						
AKO90997	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
AKO90997	Ota, T., Suzuki, Y., Nishikawa, T., Otsubi, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayaishi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahori, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, K., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kanthara, K., Katsura, N., Sato, K., Tanikawa, M., Yamazaki, M., Nishimura, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoa, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Hootera, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A., Saeki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Teraashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kunigata, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujitani, T., Oho, T., Yamada, K., Fujii, Y., Ozaki, K., Hiroo, M., Omori, Y., Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,						

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 NUCUCHI, S., ITOH, T., SHIGETA, K., SENBA, T., MATSUMURA, K.,
 NAKAJIMA, Y., MIZUNO, T., MORINAGA, M., SASAKI, M., TOGASHI, T.,
 OYAMA, M., HATA, H., WATANABE, M., KONATSU, T., MIZUSHIMA-SUGANO, J.,
 SATOH, T., SHIRAI, Y., TAKAHASHI, Y., NAKAGAWA, K., OKUMURA, K.,
 NAGASE, T., NOMURA, N., KIKUCHI, H., MASUHO, Y., YAMASHITA, R.,
 NAKAI, K., YADA, T., NAKAMURA, Y., OHARA, O., ISOGAI, T. and SUGANO, S.
 Complete sequencing and characterization of 21,243 full-length
 human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 14702039
 2
 KAWAKAMI, B., SUGIYAMA, A., TAKEMOTO, M., SUGIYAMA, T., IRIE, R.,
 OHSUKI, T., SATO, H., OGA, T., MAKIMATSU, A., ISHII, S., YAMAMOTO, J.,
 ISENO, Y., KAWAI-HIO, Y., SAITO, K., NISHIKAWA, T., KIMURA, K.,
 YAMASHITA, H., MATSUO, K., NAKAMURA, Y., SEKINE, M., KIKUCHI, H.,
 KANDA, K., WAGATA, M., MURAKAWA, K., KANEHORI, K.,
 TAKAHASHI, F., FUJITA, A., OEHIMA, A., SUZUKI, Y., SUGANO, S., NAGAHARI, K.,
 MASUHO, Y., NAGAI, K. and ISOGAI, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 2092)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
 Kazues-Kamatari, Kisarazu, Chiba 232-0812, Japan
 E-mail: genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB) ; cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1. 2092
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAMH2002191"
 /tissue_type="brain"
 /clone_id="BRAMH2"
 /note="Cloning vector: pME18SFL3"
 394..1803
 /codon_start=1
 /note="unnamed protein product"
 /protein_id="BAC03565.1"
 /db_xref="GI:21749269"
 /translation="MLKTHPKGVSEEDCLYLNIYAPAHADTSKLPVLVWPGCAFK
 TGSATIFDGSALAVIEDVIVVVOYRIGITGFPTTWDQAHPGWAFEDVOVALSWOK
 NIEFGDPSSTVITFGESAGATSVSSLIISPMKGLFHKAIMSGVALIPLBAHYE
 KSELDQVAFHCNNADSEBALRCLRTKSKELTLGSKTSFTVRVDGAFPNNEPL
 DLSOKRFAKIPSTIGVNSHGGCLLPKMEKAPILSGSNSTSLAHILIONLHLPQVL
 HUYANERFHDKHSILTEIRBSLDLLDGFVFPVVALITAYRHPAGAGLVYEEYERHPO
 GFEDTQPAFVKADHAEVRFVFGAGPLKEDIWFBEGATEBSKLSKMKMYATTPART
 CNEPNDGLSLMPAYNLTEQYLTQDILNLSIQRLKEPRVEWTSITPILIIASMSLHSP
 LLSITFLPSILQPEFFFCAP"

D	258	CCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATGCGCTGGGATTAACCTTG	317
Q	271	CGAAGAGCCACTCCCTAACCCTAATTT-----GTGCTCCAGAACTCAGAGTGGCT	320
D	318	CGAAGAGCCACTCCCTAACCCTAATTTGTAAGACGAGGCTCCCGAGAACTCAGAGTGGCT	377
Q	321	GCTCTTAAGATCAACATGCTTCAGAGTGCATTAACCGAAATTCGGAAGTGCAGAGACTG	380
D	378	GCTCTTAAGATCAACATGCTTCAGAGTGCATTAACCGAAATTCGGAAGTGCAGAGACTG	437
Q	381	CCTCAACCTGAACATCTATGCGCCCTGCGCCAGCGGATTCAGAGTCCAACTCCCGCTT	440
D	438	CCTCAACCTGAACATCTATGCGCCCTGCGCCAGCGGATTCAGAGTCCAACTCCCGCTT	497
Q	441	GATGTGATTTCCAGAGAGTGGCTTCAGAGCTGAGCTCAGCTCCATCTTTATAGGTCCGC	500
D	498	GATGTGATTTCCAGAGAGTGGCTTCAGAGCTGAGCTCAGCTCCATCTTTATAGGTCCGC	557
Q	501	CTTGGCTGCTTATAGAGAGCTGCTGTTGTGCTGCAGTACCGGCTAGAAATATTTGG	560
D	558	CTTGGCTGCTTATAGAGAGCTGCTGTTGTGCTGCAGTACCGGCTAGAAATATTTGG	617
Q	561	TTTCTTACCAATGAGATGAGATGCTCGGGGAAATGGGCTTCAGAGCAAGTGGC	620
D	618	TTTCTTACCAATGAGATGAGATGCTCGGGGAAATGGGCTTCAGAGCAAGTGGC	677
Q	621	TGCTGTGCTGCTGGGTCAGAAAGAAATCGAGTTCTTCGATGGGAGCCCAAGCTCTGTAC	680
D	678	TGCTGTGCTGCTGGGTCAGAAAGAAATCGAGTTCTTCGATGGGAGCCCAAGCTCTGTAC	737
Q	681	CATCTTTGGGAGATCGCGGGAGCCATAGTGTTCATAGCTTATACGTCTCCATGGC	740
D	738	CATCTTTGGGAGATCGCGGGAGCCATAGTGTTCATAGCTTATACGTCTCCATGGC	797
Q	741	CAAAAGCTTATTCACAAAGCCATCATGAGAGTGGGATGSCATCATCTTACCTGGA	800
D	798	CAAAAGCTTATTCACAAAGCCATCATGAGAGTGGGATGSCATCATCTTACCTGGA	857
Q	801	GGCCCATATTAATGAGAAAGTGAAGACCTGCGAGGTGTGTCACATTTCTGTGTAA	860
D	858	GGCCCATATTAATGAGAAAGTGAAGACCTGCGAGGTGTGTCACATTTCTGTGTAA	917
Q	861	TGCGTGAAGCTGAGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGCTCT	920
D	918	TGCGTGAAGCTGAGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGCTCT	977
Q	921	GACCTCAGCCGAGAAACAAAGTCTTTCACTGAGTGTGTAAGTGTCTTTCTTAA	980
D	978	GACCTCAGCCGAGAAACAAAGTCTTTCACTGAGTGTGTAAGTGTCTTTCTTAA	1037
Q	981	TGAGCCTCTAGATCTATTTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGT	1040
D	1038	TGAGCCTCTAGATCTATTTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGT	1097
Q	1041	CAATTAACAGAGAGTGGCTCTCTGCTGCTATGAGAGAGGCTCTGAGATCTCAGTGG	1100
D	1098	CAATTAACAGAGAGTGGCTCTCTGCTGCTATGAGAGAGGCTCTGAGATCTCAGTGG	1157
Q	1101	CTCCAAACAAGTCCCTGCGCTCCATCTGATTCAGAAACATCTGCAATCCCGCTCAGTA	1160
D	1158	CTCCAAACAAGTCCCTGCGCTCCATCTGATTCAGAAACATCTGCAATCCCGCTCAGTA	1217
Q	1161	TTTGACACCTTGTGCTATGATTAATCTTCATGACAAAGCACTCCCTGACTAAATTCGAGA	1220
D	1218	TTTGACACCTTGTGCTATGATTAATCTTCATGACAAAGCACTCCCTGACTAAATTCGAGA	1277
Q	1221	CAGCTCTTGAAGCTGCTGAGAGTGTGTTCTTTGTGCTCTGCACTGATCAAGCTCG	1280
D	1278	CAGCTCTTGAAGCTGCTGAGAGTGTGTTCTTTGTGCTCTGCACTGATCAAGCTCG	1337
Q	1281	ATAATCAGAGAGTGTGCTGCACTGCTCTATCTTATGAGTTTCGGAACCGGCTCAGTG	1340
D	1338	ATAATCAGAGAGTGTGCTGCACTGCTCTATCTTATGAGTTTCGGAACCGGCTCAGTG	1397

Q	1341	CTTTGAAGACAGAAAGCGGGCTTTTGTCAAAACCGACCAAGCTGATGAAATCCGCTTGT	1400
D	1398	CTTTGAAGACAGAAAGCGGGCTTTTGTCAAAACCGACCAAGCTGATGAAATCCGCTTGT	1457
Q	1401	GTTCCGCTGCTCTCTCTGAGAGGGGAGCATTTGTTATGTTTCAGAGAGCAGAGAGAGA	1460
D	1458	GTTCCGCTGCTCTCTCTGAGAGGGGAGCATTTGTTATGTTTCAGAGAGCAGAGAGAGA	1517
Q	1461	GAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTTGTCTGAAACCGGAAATCC	1520
D	1518	GAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTTGTCTGAAACCGGAAATCC	1577
Q	1521	TAAATGGAACGACCTGTCTGTGAGCCAGCTTATATATGATGATGAGAGTACCTCCAGCT	1580
D	1578	TAAATGGAACGACCTGTCTGTGAGCCAGCTTATATATGATGATGAGAGTACCTCCAGCT	1637
Q	1581	GGAATTTGAACATGAGCTTGGGACAGAGACTCAAGAAACCGGAGTGAATTTTGAACAG	1640
D	1638	GGAATTTGAACATGAGCTTGGGACAGAGACTCAAGAAACCGGAGTGAATTTTGAACAG	1697
Q	1641	CACCATCCCTGATACCTGTCTGCTCCGACATGCTCCAGATCTCTTCTTCTTAACT	1700
D	1698	CACCATCCCTGATACCTGTCTGCTCCGACATGCTCCAGATCTCTTCTTCTTAACT	1757
Q	1701	TTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCTCTTGA	1746
D	1758	TTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCTCTTGA	1803

RESULT 9
 BC039073 2298 bp mRNA linear PRI 18-NOV-2002
 LOCUS
 DEFINITION Homo sapiens, similar to carboxylesterase, clone IMAGE:5174937,
 mRNA.
 VERSION BC039073
 ACCESSION BC039073.1 GI:25058132
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (01-NOV-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 Contact: MGC help desk
 Email: cgabrs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 51 Row: K Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 Location/Qualifiers
 1..2298
 /organism="Homo sapiens"

FEATURES
 source

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174937"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"

```

ORIGIN

Query Match 86.4%; Score 1509.2; DB 9; Length 2298;
 Beel Local Similarity 92.4%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 3; Indels 132; Gaps 1;

```

QY 91 GGGCCCTGCTGGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGGCAAGCA 150
DB 103 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGGCAAGCA 162
QY 151 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTGTGCT 210
DB 163 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTGTGCT 222
QY 211 CCCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 270
DB 223 CCCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 282
QY 271 CGAGAGGCACTCCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTAAGT 330
DB 283 CGAGAGGCACTCCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTAAGT 342
QY 331 CACACATGCTCAGAGTGTATCCCGAAATTCCGAGTGTCCAGAGACTGCTCTTAAGT 390
DB 343 CACACATGCTCAGAGTGTATCCCGAAATTCCGAGTGTCCAGAGACTGCTCTTAAGT 402
QY 391 AACATCTATGCGCTGCGCCGACGCGATACAGGCTCCAGCTCCCGTCTTGTTGTTTC 450
DB 403 AACATCTATGCGCTGCGCCGACGCGATACAGGCTCCAGCTCCCGTCTTGTTGTTTC 462
QY 451 CCAGAGGAGGCTTGAAGCTGCTCAGCTGCATCTTTGATGGGTCGCGCTGCTGCTG 510
DB 463 CCAGAGGAGGCTTGAAGCTGCTCAGCTGCATCTTTGATGGGTCGCGCTGCTGCTG 522
QY 511 TATGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 569
DB 523 TATGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 582
QY 570 ----- 569
DB 583 TGAGTCTCTCAATGATGACATCCAGATCCCAAGAGGCGAGTGCAGTAGCCATA 642
QY 570 ----- 569
DB 643 GACAGTCCCGACCACTTAAGAGAGCGAGCTTCCGGGACAAATGAACAAAATCGGCC 702
QY 570 ----- 569
DB 703 TAACTTACCTAGACATGAGATGATGCTCCGGGAACTGAGGCTTCAAGACAGAGTG 762
QY 619 GCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678
DB 763 GCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822
QY 679 AACATCTTGGGAGATCCGCGGAGCCAAATGTTTCAAGCTTAACTGCTCCCAAG 738
DB 823 AACATCTTGGGAGATCCGCGGAGCCAAATGTTTCAAGCTTAACTGCTCCCAAG 882
QY 739 GCCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGTGCATCATCCCTTAAGCTG 798
DB 883 GCCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGTGCATCATCCCTTAAGCTG 942
QY 799 GAGGCCCATGATTTATGAGAGAGTGAAGACTGCGAGGTGCTTGCACATTTCTGTGTAAC 858
DB 943 GAGGCCCATGATTTATGAGAGAGTGAAGACTGCGAGGTGCTTGCACATTTCTGTGTAAC 1002

```

```

QY 859 AATGCGTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGGACAAAACCTTCAAGAGCTG 918
DB 1003 AATGCGTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGGACAAAACCTTCAAGAGCTG 1062
QY 919 CTGACCTTCAGCCAGAGAAAACAAAGTCTTTCACTGAGAGTGTGATGCTTTCTTCT 978
DB 1063 CTGACCTTCAGCCAGAGAAAACAAAGTCTTTCACTGAGAGTGTGATGCTTTCTTCT 1122
QY 979 AATGAGCCCTTAAGATCTAATGCTCAGAAAGCATTTAAAGCAATCTTCCATCATCGGA 1038
DB 1123 AATGAGCCCTTAAGATCTAATGCTCAGAAAGCATTTAAAGCAATCTTCCATCATCGGA 1182
QY 1039 GTCAATTAACACAGAGTGTGCTTCTGCTGCTTATGAAGAGGCTCTGAGATCCCTCACT 1098
DB 1183 GTCAATTAACACAGAGTGTGCTTCTGCTGCTTATGAAGAGGCTCTGAGATCCCTCACT 1242
QY 1099 GGGTCCAAAGAGTCCCTTGCCTCATCTGATCAAAAATCTTGCACATCCGCTCAG 1158
DB 1243 GGGTCCAAAGAGTCCCTTGCCTCATCTGATCAAAAATCTTGCACATCCGCTCAG 1302
QY 1159 TATTTGCACTTTGAGGCTTAATGAATACCTTCAATGAACAAGCACTCCGATGAATCCGA 1218
DB 1303 TATTTGCACTTTGAGGCTTAATGAATACCTTCAATGAACAAGCACTCCGATGAATCCGA 1362
QY 1219 GACAGCTTTCTGAGACTTGTGCTGAGATGTTCTTTGTGCTGCTGCACTGATCAAGCT 1278
DB 1363 GACAGCTTTCTGAGACTTGTGCTGAGATGTTCTTTGTGCTGCTGCACTGATCAAGCT 1422
QY 1279 CGATATCAAGAGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1338
DB 1423 CGATATCAAGAGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1482
QY 1339 TGCTTTGAAGACAGCAAGCCGCTTTGTCAAAGCCGACACGCTGATGAAGTCCGCTTT 1398
DB 1483 TGCTTTGAAGACAGCAAGCCGCTTTGTCAAAGCCGACACGCTGATGAAGTCCGCTTT 1542
QY 1389 GTGTTGCGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1458
DB 1543 GTGTTGCGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1602
QY 1459 GAGAACTTACTGAGCCGGAAGATGTAATCTGAGCTTACTTGTGCTGAAACCGGAAT 1518
DB 1603 GAGAACTTACTGAGCCGGAAGATGTAATCTGAGCTTACTTGTGCTGAAACCGGAAT 1662
QY 1519 CCTAATGGAAGACACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1578
DB 1663 CCTAATGGAAGACACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722
QY 1579 CTGACCTTGAACATGAGCTCGGACAGAGACTCAAGAGACCGGAGTGAATTTTGAC 1638
DB 1723 CTGACCTTGAACATGAGCTCGGACAGAGACTCAAGAGACCGGAGTGAATTTTGAC 1782
QY 1639 AGCACCATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1698
DB 1783 AGCACCATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1842
QY 1699 ACTTTCCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
DB 1843 ACTTTCCTCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1890

```

RESULT 10
 CQ733044 1716 bp DNA linear PAT 03-FEB-2004
 LOCUS CQ733044
 DEFINITION Sequence 18978 from Patent WO02068579.
 ACCSSION CQ733044
 VERSION CQ733044.1 GI:42314969
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1

AUTHORS

Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

JOURNAL

Patent: WO 02068579-A 18978 06-SEP-2002;

FEATURES

source

1. 1716
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity 91.6%; Score 1425; DB 6; Length 1716;
Matches 1428; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 84 AAAAAGGAGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTCAGG 143
DB 21 AAAAGAGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTCAGG 80
QY 144 CAGCAAGTCACTGTCTGGAGAGCCCTGCTGCTGAAAGTTCCTGAGAGTCCCTT 203
DB 81 CAGCAAGTCACTGTCTGGAGAGCCCTGCTGCTGAAAGTTCCTGAGAGTCCCTT 140
QY 204 TGTCTCTCCCGCTGGAGATCCCTGCTGATTAAGAACCCGACGCTGATCCCTGGGA 263
DB 141 TGTCTCTCCCGCTGGAGATCCCTGCTGATTAAGAACCCGACGCTGATCCCTGGGA 200
QY 264 TAACCTGAGAGAGCCACCTCCCTAATTTGAGCTCCAGAACTCAGATGCTGCT 323
DB 201 TAACCTGAGAGAGCCACCTCCCTAATTTGAGCTCCAGAACTCAGATGCTGCT 260
QY 324 CTTAAGTCAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTGAGAGCTGCT 383
DB 261 CTTAAGTCAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTGAGAGCTGCT 320
QY 384 CTACCTGAACATCTATGAGCTCCGACGCGGATACAGGCTTCAAGCTCCCTGCT 443
DB 321 CTACCTGAACATCTATGAGCTCCGACGCGGATACAGGCTTCAAGCTCCCTGCT 380
QY 444 GTGCTCCAGAGAGGCTCTTCAAGCTGCTCAGCTCCATCTTGTGAGGCTCCGCT 503
DB 381 GTGCTCCAGAGAGGCTCTTCAAGCTGCTCAGCTCCATCTTGTGAGGCTCCGCT 440
QY 504 GGCTGCTATAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563
DB 441 GGCTGCTATAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500
QY 564 CTTCACACATGAGATCAGCATGCTCGGGGAACTGGGCTTCAAGAGCCAGGCTGCT 623
DB 501 CTTCACACATGAGATCAGCATGCTCGGGGAACTGGGCTTCAAGAGCCAGGCTGCT 560
QY 624 TCTGCTCTGGGCTCAAGAAACATGATTTCTGCTGGAGAGCCAGCTGTGACCAT 683
DB 561 TCTGCTCTGGGCTCAAGAAACATGATTTCTGCTGGAGAGCCAGCTGTGACCAT 620
QY 684 CTTTGGGAGGCTCGGGGAGCCATTAAGTCTTCAAGCTTATTAAGTCTTCCATGCT 743
DB 621 CTTTGGGAGGCTCGGGGAGCCATTAAGTCTTCAAGCTTATTAAGTCTTCCATGCT 680
QY 744 AGGCTTATTCACAAAGCCATCATGAGAGTGGGCTGCTGCTGCTGCTGCTGCT 803
DB 681 AGGCTTATTCACAAAGCCATCATGAGAGTGGGCTGCTGCTGCTGCTGCTGCT 740
QY 804 CCAATGATTAAGAGAGAGTGAAGCTGCAAGTGGTTCATTTCTGCTGATTAAGATG 863
DB 741 CCAATGATTAAGAGAGAGTGAAGCTGCAAGTGGTTCATTTCTGCTGATTAAGATG 800
QY 864 GTGAGACTGAGAGGCTGCTGAGAGTGGCTGAGAGCAAAACCTTCAAGAGAGCTGCT 923
DB 801 GTGAGACTGAGAGGCTGCTGAGAGTGGCTGAGAGCAAAACCTTCAAGAGAGCTGCT 860
```

```
QY 924 CCTCAGCCAGAAACAAAGCTTTTCACTGAGTGTGATGAGTGTCTTTCTTAATGA 983
DB 861 CCTCAGCCAGAAACAAAGCTTTTCACTGAGTGTGATGAGTGTCTTTCTTAATGA 920
QY 984 GCTCTAAGTCTATGCTGCTGAGAGAGCTTTAAAGCAATTCCTTCATCATGAGTCA 1043
DB 921 GCTCTAAGTCTATGCTGCTGAGAGAGCTTTAAAGCAATTCCTTCATCATGAGTCA 980
QY 1044 TAACCAAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
DB 981 TAACCAAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
QY 1104 CAGCAAGTCCCTTGGCTTCATCTGATTAAGAAACATCTGCAATCCGCTCAGTATT 1163
DB 1041 CAGCAAGTCCCTTGGCTTCATCTGATTAAGAAACATCTGCAATCCGCTCAGTATT 1100
QY 1164 GCACCTTGTGCTATGATTAATCTTCAATGACAAACATCCCTGCTGCTGCTGCTGCT 1223
DB 1101 GCACCTTGTGCTATGATTAATCTTCAATGACAAACATCCCTGCTGCTGCTGCTGCT 1160
QY 1224 TCTTCTGAGCTTCTGAGAGTGTCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1283
DB 1161 TCTTCTGAGCTTCTGAGAGTGTCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
QY 1284 TCACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
DB 1221 TCACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280
QY 1344 TGAAGACAGAGAGGCTTCTTGTCAAGGCGCAGCAGCTGATGAGTCCGCTTGTGCT 1403
DB 1281 TGAAGACAGAGAGGCTTCTTGTCAAGGCGCAGCAGCTGATGAGTCCGCTTGTGCT 1340
QY 1404 CGGTGCTGCTTCTGTAAGGCGGAGCATTTTATGTTGTAAGAGCAACGAGAGAGAA 1463
DB 1341 CGGTGCTGCTTCTGTAAGGCGGAGCATTTTATGTTGTAAGAGCAACGAGAGAGAA 1400
QY 1464 GTTACTGAGCGGAGAGTGAATTAATGAGGCTTACCTTGTGCTGCAACGGGGA 1516
DB 1401 GTTACTGAGCGGAGAGTGAATTAATGAGGCTTACCTTGTGCTGCAACGGGGA 1453
```

RESULT 11
AX958392 1857 bp DNA linear PAT 14-JAN-2004
LOCUS AX958392
DEFINITION Sequence 20 from Patent WO0226988.
ACCESSION AX958392
VERSION AX958392.1 GI:40879350
KEYWORDS

SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Human drug metabolizing enzymes

AUTHORS

Patent: WO 0226988-A 20 04-Apr-2002;

JOURNAL

Incyte Genomics, Inc. (US)

FEATURES

Location/Qualifiers
1..1857

source

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 79.4%; Score 1386.2; DB 6; Length 1857;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

```
QY 81 GGGAAAACTGGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTC 140
DB 264 GGTCAAGAGAGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTC 323
QY 141 GGGCAAGCAAGTCACTGTCTGGAGAGCCCTGCTGTAAGTGTCTCTGAGAGTCC 200
```

```

Db      324 GGGGAGGAGGACATGCTGCTGGAGAGCCCTGCTGGAAGCGTTCCTCGAGTCCC 383
Qy      201 CTTTCTGCTGCTCCCGCTGGAGATCCCTGCGATTTTAAGAACCCGCGAGCTGCAATCCGCTCG 260
Db      384 CTTTCTGCTGCTCCCGCTGGAGATCCCTGCGATTTTAAGAACCCGCGAGCTGCAATCCGCTCG 443
Qy      261 GGAATTAATCTGGAGAGGACCTCCCTAACCCTAATTTGTCCTCGAAGACTGAGAGTGT 320
Db      444 GGAATTAATCTGGAGAGGACCTCCCTAACCCTAATTTGTCCTCGAAGACTGAGAGTGT 503
Qy      321 GCTCTTAAGATCAACAATGCTCAAGGTGCAATTAACCGAAATTCGAGTGTCAAGAACTG 380
Db      504 GCTCTTAAGATCAACAATGCTCAAGGTGCAATTAACCGAAATTCGAGTGTCAAGAACTG 563
Qy      381 CCTCACTCACTGAACATCTATGCGCCCTGCCAGCCGATACAGAGCTCCAGCTCCGCTCT 440
Db      564 CCTCACTCACTGAACATCTATGCGCCCTGCCAGCCGATACAGAGCTCCAGCTCCGCTCT 623
Qy      441 GGTGTGTTCCAGAGAGGCTCTCAAGAATGAGCTCAGGCTCACTTTGATGGTCCGC 500
Db      624 GGTGTGTTCCAGAGAGGCTCTCAAGAATGAGCTCAGGCTCACTTTGATGGTCCGC 683
Qy      501 CCTGCTGCTGATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
Db      684 CCTGCTGCTGATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
Qy      561 TTTCTTCAACACATGAGATCAGCATGCTCCGGGAACTGCGGCTTCAAGAACAGAGTGC 620
Db      744 TTTCTTCAACACATGAGATCAGCATGCTCCGGGAACTGCGGCTTCAAGAACAGAGTGC 803
Qy      621 TGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680
Db      804 TGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 863
Qy      681 CATCTTTGGCGAGTCCGCGGAGCATAAGTCTTTCTAATCTTATCTGCTCCCATGAGC 740
Db      864 CATCTTTGGCGAGTCCGCGGAGCATAAGTCTTTCTAATCTTATCTGCTCCCATGAGC 923
Qy      741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGCATCATCTCTTACCTGGA 800
Db      924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGCATCATCTCTTACCTGGA 983
Qy      801 GGGCCATGATTAAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 860
Db      984 GGGCCATGATTAAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1043
Qy      861 TGGCTCAGACTCTGAGAGCTGCTGAGAGTGCCTGAGAGCAAAACCTCCAGAGAGTGT 920
Db      1044 TGGCTCAGACTCTGAGAGCTGCTGAGAGTGCCTGAGAGCAAAACCTCCAGAGAGTGT 1103
Qy      921 GACCTTCAGCCGAGAAAACAAAGTCTTTCACTGAGTGTGTTGAGTGTCTTTCTTAA 980
Db      1104 GACCTTCAGCCGAGAAAACAAAGTCTTTCACTGAGTGTGTTGAGTGTCTTTCTTAA 1163
Qy      981 TGAAGCTCTAAGTCTAATGCTCTCAGAAAGCAATTAAGCAATTCCTCCATCATGAGAGT 1040
Db      1164 TGAAGCTCTAAGTCTAATGCTCTCAGAAAGCAATTAAGCAATTCCTCCATCATGAGAGT 1223
Qy      1041 CAATTAACAGAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
Db      1224 CAATTAACAGAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
Qy      1101 CTCACAAAGTCCCTTGCCTTCATCTGATACAAAACATCTGCAACATCCGCTCAGTA 1160
Db      1256 CTCACAAAGTCCCTTGCCTTCATCTGATACAAAACATCTGCAACATCCGCTCAGTA 1274
Qy      1161 TTTGACCTTGTGCTAATGATATCTTCATGACAGAGCTCCCTGAGCTGAATCCGAGA 1220
Db      1275 TTTGACCTTGTGCTAATGATATCTTCATGACAGAGCTCCCTGAGCTGAATCCGAGA 1334
Qy      1221 CAGTCTTGTGAGCTTGTGAGATGTGTTCTTTGTGTGCTCCCTGACATGATCAGAGCTG 1280

```

```

Db      1335 CAGTCTTGTGAGCTTGTGAGAGATGTGTTCTTTGTGTGCTCCCTGACATGATCAGCTG 1394
Qy      1281 ATATCAGAGAGATGTGTGAGACCTGTCTAATCTTATGAGTGTGGGACCCGCTCAGTG 1340
Db      1395 ATATCAGAGAGATGTGTGAGACCTGTCTAATCTTATGAGTGTGGGACCCGCTCAGTG 1454
Qy      1341 CTTTGAGACAGAGAGCCGCTTTTGTCAAGCCGACCAAGCTGATGAGTCCGCTTGT 1400
Db      1455 CTTTGAGACAGAGAGCCGCTTTTGTCAAGCCGACCAAGCTGATGAGTCCGCTTGT 1514
Qy      1401 GTTCGCTGTGCTCTTCTGAGAGGGGACATTTGTTATGTTCCAGAGAGCCAGAGAGAG 1460
Db      1515 GTTCGCTGTGCTCTTCTGAGAGGGGACATTTGTTATGTTCCAGAGAGCCAGAGAGAG 1574
Qy      1461 GAGTCTTGAAGCCGAGAGATGATGAATPACTGGGCTACTTGTGCTCGAACCGGAGATCC 1520
Db      1575 GAGTCTTGAAGCCGAGAGATGATGAATPACTGGGCTACTTGTGCTCGAACCGGAGATCC 1634
Qy      1521 TAATGGAGAGCAGCTGTCTGTGCGCAGCTTATATCTGACTGAGACATCTCAGCT 1580
Db      1635 TAATGGAGAGCAGCTGTCTGTGCGCAGCTTATATCTGACTGAGACATCTCAGCT 1694
Qy      1581 GGACTTGAACATGAGCTCTCGAGAGACTCAAGAGACCCGCGAGAGATGTGTG 1635
Db      1695 GGACTTGAACATGAGCTCTCGAGAGACTCAAGAGACCCGCGAGAGATGTGTG 1749

```

```

RESULT 12
BC069548      1896 bp      mRNA      linear      PRI 30-JUN-2004
LOCUS      BC069548
DEFINITION      Homo sapiens hypothetical protein FLJ31547, mRNA (CDNA clone
ACCESSION      MGC:97193 IMAGE:7262439), complete cds.
VERSION      BC069548
KEYWORDS      BC069548.1 GI:46854430
SOURCE      MGC.
ORGANISM      Homo sapiens (human)
             Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1896)
AUTHORS      Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
             Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
             Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
             Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hathe,F.,
             Diatchenko,L., Marusik,K., Farmer,A.A., Rubin,G.M., Hong,L.,
             Scapleorn,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
             Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,
             Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,
             McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
             Wortley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
             Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
             Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
             Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shenchenko,Y.,
             Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
             Dickson,M.C., Rodriguez,A.C., Gilwood,J., Schmitt,J., Myers,R.M.,
             Buterfield,Y.S., Kozminski,M.I., Skalske,U., Smilans,D.E.,
             Scherch,A., Schein,J.B., Jones,S.J., and Marra,M.A.
             Generation and initial analysis of more than 15,000 full-length
             human and mouse cDNA sequences
             Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL      2 (bases 1 to 1896)
PUBMED      Strausberg,R.
AUTHORS      Direct Submission
TITLE      Submitted (29-APR-2004) National Institutes of Health, Mammalian
JOURNAL      Gene Collection (MGC), Cancer Genomics Office, National Cancer
             Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
             USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
             Contact: MGC help desk
             Email: gcgpbp-remail.nih.gov
             Tissue Procurement: Baylor Human Genome Sequencing Center

```


CDNA Library Preparation: Baylor Human Genome Sequencing Center
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loubeled, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambatti,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ULNL at: <http://image.jnl.gov>
Series: IRBR Plate: 4 Row: f Column: 1.

FEATURES

source

1..1896

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:97193 IMAGE:7262439"

/issue_type="PCR rescued clones"

/clone_id="NIH MGC 244"

/note="Vector: PCR-Script Amp SK(+)"

1..1896

/gene="FLJ31547"

/db_xref="LOCUSID:221223"

120..1697

/gene="FLJ31547"

/codon_start=1

/product="hypothetical protein FLJ31547"

/protein_id="AAH69548.1"

/db_xref="GI:46854431"

/translation="MSGNVHPGQILIMVLAATKGPSAEPQORNTRLGMOCKO
VTVLSPVYVNFVLPVFAAPPLGSLRFTNPQASPMNLRATSPYLQNSBML
LDQHLKHYPKFVSEBDLYNTPAHADGSLPLVWFGPAFTGSAISIDGS
ALAAVEDVAVVVOYRIGIFGFETTDQAPGNAPKQVVALSVQNLIFFGDPS
SVTIRGSAALSVSLLISPMKGLFHALMESGVALIPLYEADYKSDLDQVAF
FGNNAISDEALRLCLRTKPSKELTLQKTKSPFVVDGAFPMEDLSOKAFK
IPSTIGVNNHCEGLFPMKEAPEILSGNKSIALHILNIHLIPIQVLAHYNEFH
KSLTEIRLSDLDLDGDFEVVPAITARYHREGATEBKILSRMMQVMAVTPARD
PNGNDLSMPAYNLTEOYLQDLNMSLQRLKEPRVDVWTSTIPLISASDMHSPLS
SLTFLSLQPPFFCAP"

ORIGIN

Query Match 77.0%; Score 1344.4; DB 9; Length 1896;
Best Local Similarity 90.9%; Pred. No. 0;

Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

91 GGGCTTCTGCTGAAGGGCCACAGAGGAACACCGAGCTGGATGATTCAAGGCAAGCA 150
192 GGGCTTCTGCTGAAGGGCCACAGAGGAACACCGAGCTGGATGATTCAAGGCAAGCA 251
151 GTCACTGTGCTGGAAGGCTGTGCTGTGAACGTGTCCTCGAGATCCCTTTGCTGCT 210
252 GTCACTGTGCTGGAAGGCTGTGCTGTGAACGTGTCCTCGAGATCCCTTTGCTGCT 311
211 CCCCCGCTGGATCCCTCGCATTTACGAACCCGAGCCTGCATGCGCTTGGATTAATTG 270
312 CCCCCGCTGGATCCCTCGCATTTACGAACCCGAGCCTGCATGCGCTTGGATTAATTG 371
271 CGAAGAACCACTCTCACTCCCTAATTTTGCTGCTCGAAGCTCAAGTGGCTGCTTATGAT 330
372 CGAAGAACCACTCTCACTCCCTAATTTTGCTGCTCGAAGCTCAAGTGGCTGCTTATGAT 431
331 CAACACATGCTCAAGGTGATTAACCGAAATTGGAGTGTCAAGAGACTGCTTACCTG 390
432 CAACACATGCTCAAGGTGATTAACCGAAATTGGAGTGTCAAGAGACTGCTTACCTG 491
391 AAGATTAATGCTGCTGCAAGCGCGATACAGGCTCCAGCTCCCGCTTGTGTGCTTC 450
492 AAGATTAATGCTGCTGCAAGCGCGATACAGGCTCCAGCTCCCGCTTGTGTGCTTC 551

451 CCAGAGGTGCTTCAAGACTGAGCTCAGCTCCATCTTTGATGGTCCGCGCTGCTGCC 510
552 CCAGAGGTGCTTCAAGACTGAGCTCAGCTCCATCTTTGATGGTCCGCGCTGCTGCC 611
511 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
612 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
571 ACATGGATCAGATGCTCCGCGGGAATCGGCTTCAAGAGCTGAGTGGCTGCTGCTGCC 630
672 ACATGGATCAGATGCTCCGCGGGAATCGGCTTCAAGAGCTGAGTGGCTGCTGCTGCC 731
631 TGGGTCCAGAGAACATGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 690
732 TGGGTCCAGAGAACATGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
691 GAGTCCGCGGAGCATGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
792 GAGTCCGCGGAGCATGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
751 TTCCACAAAGCATCATGAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 810
852 TTCCACAAAGCATCATGAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911
811 TATGAGAGAGTGAAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
912 TATGAGAGAGTGAAGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
871 TCTGAGGCTGCTGAGAGTGGCTGAGAGCAAAACCCGCAAGAGAGCTGAGACCTTCA 930
972 TCTGAGGCTGCTGAGAGTGGCTGAGAGCAAAACCCGCAAGAGAGCTGAGACCTTCA 1031
931 CAGAAACAAAGCTTCTCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 990
1032 CAGAAACAAAGCTTCTCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1091
991 GATCTATGCTGCTGAGAGATTTAAAGCAATTCCTTCCATCATCGAGTCAATTAACAC 1050
1092 GATCTATGCTGCTGAGAGATTTAAAGCAATTCCTTCCATCATCGAGTCAATTAACAC 1151
1051 GAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
1152 GAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
1111 TCCCTTCCCTCCATGATGATTAACAAACATCTGCAATCCGCTCCATGATTTGCACTT 1170
1212 TCCCTTCCCTCCATGATGATTAACAAACATCTGCAATCCGCTCCATGATTTGCACTT 1271
1171 GTGGCTAATGATTAATCTTCAATGACAGCACTCCCTGCACTGAATCCGAGACAGTCTTC 1230
1272 GTGGCTAATGATTAATCTTCAATGACAGCACTCCCTGCACTGAATCCGAGACAGTCTTC 1331
1231 GACTTCTTGGAGATGCTGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290
1332 GACTTCTTGGAGATGCTGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1391
1291 GATGCTGAGCACTGTCTAATCTATGATGTTCCGACCCGCTCAAGTCTTTGAAGAC 1350
1392 ----- 1391
1351 ACGAAGCGGCTTTTGTCAAGCCGACGACGCTGATGAAGTCCGTTGTGCTGGTGGT 1410
1392 ----- 1391
1411 GCTTCTGGAAGGGGACATGTTATGTTCAAGAGGCTGAGAGAGAGAGATTAATG 1470
1392 -----GAAAGGCAAGAGAGAGAGAGATTAATG 1421
1471 AGCCGAAAGATGATGAATTAATGAGCTTCTGCTGCAACCGGGAATCTTAATGGGAC 1530
1422 AGCCGAAAGATGATGAATTAATGAGCTTCTGCTGCAACCGGGAATCTTAATGGGAC 1481
1531 GACTGCTCTGCTGAGCCAGCTTAATCTGATGAGAGTACCTTCAAGCTGAGCTTGAAC 1590

QY 1591 ATGAGCTCGGACGAGACTCAAGAACCGGGGTGATTTTGGACACGACCATCTCCC 1650
 DB 1572 ATGAGCTCGGACGAGACTCAAGAACCGGGGTGATTTTGGACACGACCATCTCCC 1631
 QY 1651 CTGATCTGTCTGCTGCTCGACATGCTCCACAGTCTCTTTCTTCTTAACCTTCTCTCT 1710
 DB 1632 CTGATCTGTCTGCTGCTCGACATGCTCCACAGTCTCTTTCTTCTTAACCTTCTCTCT 1691
 QY 1711 CTGCTCAGGCTTCTTTCTTTCTTTGCTCTTGA 1746
 DB 1692 CTCTCAGGCTTCTTTCTTTGCTCTTGA 1727
 RESULT 14
 AK056109
 LOCUS 1962 bp mRNA linear PRI 30-JAN-2004
 DEFINITION Homo sapiens cDNA FLJ1547 fib, clone NT2R12001010, weakly similar
 to FAT1Y ACYL-COA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
 AK056109
 VERSION AK056109.1 GI:16551422
 KEYWORDS oligo capping; fib (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
 Sekine, M., Obayashi, M., Nishi, T., Shidohara, T., Tanaka, T.,
 Ishii, S., Yamamoto, K., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahara, K., Murakami, K., Yasuda, K., Iwayanagi, T., Magatsuma, M.,
 Shiratori, A., Sudo, H., Hosokawa, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
 Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
 Hirooka, K., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
 Hiroda, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,
 Hara, H., Tanabe, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuuki, H., Oshima, A.,
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
 Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
 Kunagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
 Tachiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
 Ozaki, K., Hirao, M., Ohmori, Y., Okamoto, S., Okitani, R., Kobatake, N.,
 Noguchi, S., Itoh, T., Shigeta, K., Senda, T., Matsunawa, K.,
 Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T.,
 Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
 Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
 Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T., and Sugano, S.
 Complete sequencing and characterization of 21,243 full-length
 human cDNAs
 Nac. Genet. 36 (1), 40-45 (2004)
 14702039
 2
 Nishimura, K., Magatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
 Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, B., Omura, Y., Abe, K.,
 Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamashita, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, K., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
 Kikkawa, E., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
 Kikuchi, H., Murakami, K., Kanehori, K., Takahashi, Fujii, A.,
 Oshima, A., Sugiyama, A., Kamihara, B., Suzuki, Y., Sugano, S.,
 Nagahara, K., Masuho, Y., Nagai, K., and Isegai, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 1662)

AUTHORS Isegai, T., Otsuki, T. and Sugiyama, T.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2001) Takao Isegai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yano, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.
 FEATURES
 source
 Location/Qualifiers
 1..1962
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2R12001010"
 /cell_line="NT2"
 /cell_type="teratocarcinoma"
 /clone_lib="NT2R12"
 /note="cloning vector: pMB18FL3--mRNA from NT2 neuronal
 precursor cells treated 2-weeks mitotic inhibitor after
 5-weeks retinoic acid (RA) induction..majorly NT2 neuron"
 150..1727
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB71094.1"
 /db_xref="GI:16551423"
 CDS
 ORIGIN
 Query Match 77.0%; Score 1344.4; DB 9; Length 1962;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;
 91 GGGCTTTCTGCTGAGGCGGACAGAGAAACCCAGGCTGGATGATTCAGGCAAGCA 150
 222 GGGCTTTCTGCTGAGGCGGACAGAGAAACCCAGGCTGGATGATTCAGGCAAGCA 281
 151 GTCATCTGTGCTGAGGCGGCTGCTGAGAGCTGCTGAGAGTCCCTTTGCTGCT 210
 282 GTGATCTGTGCTGAGGCGGCTGCTGAGAGCTGCTGAGAGTCCCTTTGCTGCT 341
 211 CCCCCTGCTGAGATCCCTGAGATTTAGAACCCGAGCTGATGCGCTGGATTAATCT 270
 342 CCCCCTGCTGAGATCCCTGAGATTTAGAACCCGAGCTGATGCGCTGGATTAATCT 401
 271 CGAGAGGCGGCTGAGATCCCTGAGATTTAGAACCCGAGCTGATGCGCTGGATTA 330
 402 CGAGAGGCGGCTGAGATCCCTGAGATTTAGAACCCGAGCTGATGCGCTGGATTA 461
 331 CAACATGCTCAAGATGATTAACCGAATTCGAGAGTGAAGAGTGAAGTCTCACTG 390
 462 CAACATGCTCAAGATGATTAACCGAATTCGAGAGTGAAGAGTGAAGTCTCACTG 521
 391 AACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
 522 AACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
 451 CGAGAGGCGGCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTG 510
 582 CGAGAGGCGGCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTG 641

QY 511 TATGAGACGTCGTGTTGTCGTCAGTACCCGCTAGGAATATTTGGTTTCTTAC 570
 DB 642 TATGAGACGTCGTGTTGTCGTCAGTACCCGCTAGGAATATTTGGTTTCTTAC 701
 QY 571 ACATGGGATCAGATGCTCCGGGGAATCGGGCTTCAAGACCAAGTGTCTGTCTC 630
 DB 702 ACATGGGATCAGATGCTCCGGGGAATCGGGCTTCAAGACCAAGTGTCTGTCTC 761
 QY 631 TGGGTTCAGAAACATCGAGTTCTTCCGTGGGGAACCCAGCTCTGTGACATCTTGGC 690
 DB 762 TGGGTTCAGAAACATCGAGTTCTTCCGTGGGGAACCCAGCTCTGTGACATCTTGGC 821
 QY 691 GAGTCGGGGGAGCCATAGTTTCTAGTCTTATCTGTCCTTCCATGAGCAAGGCTTA 750
 DB 822 GAGTCGGGGGAGCCATAGTTTCTAGTCTTATCTGTCCTTCCATGAGCAAGGCTTA 881
 QY 751 TTCCAAGGATCATGAGAGAGTGGGCTGATCATCTTACCTGAGGCCCATGAT 810
 DB 882 TTCCAAGGATCATGAGAGAGTGGGCTGATCATCTTACCTGAGGCCCATGAT 941
 QY 811 TATGAGAGAGTGAAGACCTGCAAGTGTGTCACATTTCTGTGTAAACAATGCTCAGAC 870
 DB 942 TATGAGAGAGTGAAGACCTGCAAGTGTGTCACATTTCTGTGTAAACAATGCTCAGAC 1001
 QY 871 TCTGAGGCGCTGTCAGTGGCTGAGGAGCAAAACCTCAAGAGAGTGTGACCCCTCAGC 930
 DB 1002 TCTGAGGCGCTGTCAGTGGCTGAGGAGCAAAACCTCAAGAGAGTGTGACCCCTCAGC 1061
 QY 931 CAGAAAACAAAGTCTTTCACCTGAGAGTGTGATGTGCTTCTTCTTAATAGAGCTCTTA 990
 DB 1062 CAGAAAACAAAGTCTTTCACCTGAGAGTGTGATGTGCTTCTTCTTAATAGAGCTCTTA 1121
 QY 991 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTTCCATCATGTGAGTCAATAACAC 1050
 DB 1122 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTTCCATCATGTGAGTCAATAACAC 1181
 QY 1051 GAGTGTGCTTCTGTCGCTTATGAAAGAGGCTCTGAGAAATCTCAGTGTCCCAACAG 1110
 DB 1182 GAGTGTGCTTCTGTCGCTTATGAAAGAGGCTCTGAGAAATCTCAGTGTCCCAACAG 1241
 QY 1111 TCCCTGCGCTCATCTGATACAAACATCTGTCACATCCGCTCAGATATTGACCTT 1170
 DB 1242 TCCCTGCGCTCATCTGATACAAACATCTGTCACATCCGCTCAGATATTGACCTT 1301
 QY 1171 GTGGCTAATGAATCTTCCATGACAAAGCATCTCCCTGATGAATCCGAGACAGTCTTCTG 1230
 DB 1302 GTGGCTAATGAATCTTCCATGACAAAGCATCTCCCTGATGAATCCGAGACAGTCTTCTG 1361
 QY 1231 GACTGTGCTGAGAGTGTCTTGTGTGTCCTGTCACATCATCAGCTGATATTCACACA 1290
 DB 1362 GACTGTGCTGAGAGTGTCTTGTGTGTCCTGTCACATCATCAGCTGATATTCACACA 1421
 QY 1291 GATGTGTGTCACCTGTCTACTTCTATGAGTTTCGGCACCGGCTCAGTCTTGAAGAC 1350
 DB 1422 ----- 1421
 QY 1351 ACGAAGCCGGCTTTGTCAAGCCGACACGCTGATGAAGTCGCTTGTGTGCTGCT 1410
 DB 1422 ----- 1421
 QY 1411 GCCTTCTGAGAGGGGACATTTGTATGTTCAAGAGCCACGAGAGAGAGAAATTAATCTG 1470
 DB 1422 ----- 1470
 QY 1471 AGCCGGAAGATGATGAATTAATCTGAGGCTTGTGCTGCAACCGGGAATCTTAATGGGAAC 1530
 DB 1452 AGCCGGAAGATGATGAATTAATCTGAGGCTTGTGCTGCAACCGGGAATCTTAATGGGAAC 1511
 QY 1531 GACCTGTCTCTGTGGCAGCTTATTAATCTGACTGAGCAGTACTCTCAGCTGGAATTTGAAC 1590
 DB 1512 GACCTGTCTCTGTGGCAGCTTATTAATCTGACTGAGCAGTACTCTCAGCTGGAATTTGAAC 1571

QY 1591 ATGAGCTTCGAGACAGACTCAAAAGACCGCGGTGAGATTTTGTGACAGACACATCCCC 1650
 DB 1572 ATGAGCTTCGAGACAGACTCAAAAGACCGCGGTGAGATTTTGTGACAGACACATCCCC 1631
 QY 1651 CTGATCTGTCTGTGCTTCGAGACATGCTCCACAGTCTCTTCTTCTTAATCTTCTCTCT 1710
 DB 1632 CTGATCTGTCTGTGCTTCGAGACATGCTCCACAGTCTCTTCTTCTTAATCTTCTCTCT 1691
 QY 1711 CTCTCCAGGCTTCTTCTTCTTGTGTGCTCTTGA 1746
 DB 1692 CTCTCCAGGCTTCTTCTTCTTGTGTGCTCTTGA 1727

RESULT 15

AB186392 1728 bp mRNA linear MM 06-AUG-2004
 LOCUS Canis familiaris cauxin mRNA for carboxylesterase-like urinary
 DEFINITION excreted protein, complete cds.

AB186392
 AB186392.1 GI:51014274

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Canis familiaris (dog)
 Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
 1 Miyazaki, M., Yamashita, T., Taira, H. and Suzuki, A.
 cauxin family protein

JOURNAL
 2 (bases 1 to 1728)
 Published Only in Database (2004)

REFERENCE
 2 (bases 1 to 1728)
 Miyazaki, M., Yamashita, T., Taira, H. and Suzuki, A.
 Direct Submission

JOURNAL
 Submitted (04-AUG-2004) Masao Miyazaki, RIKEN, Frontier System
 Research, Supra-Biomolecular System Research Group, Hiroshima 2-1,
 Wako-shi, Saitama 351-0198, Japan

(E-mail: mmiyazaki@postman.riken.jp, Tel:81-48-467-9619,
 Fax:81-48-467-9620)

FEATURES

. source

1..1728
 /location=Qualifiers
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /chromosome="2"
 1..1728

gene

1..1728
 /gene="cauxin"

CDS

1..1728
 /gene="cauxin"
 /codon_start=1
 /product="carboxylesterase-like urinary excreted protein"
 /protein_id="BBD35015.1"

/db_xref="GI:51014275"
 /translation="MSGWGHGQTLINAVVLAATGPAADAEVSTRIGWVRGQ
 ATVLGSTMVNVFLGIPEAAPPLGLRFRKRPALINDSRDPSYPLCLQNSVWL
 SDQHLKHYPNLSESDCLYNIYAAPHATGSKLPVMVFPGAFETGASIPDGS
 ALAAVEDVLIIVTQYRLGIPGPFKGDHAGNNAFLPOLALITWQENIEFGDPH
 SVTIKESGALIVSGVILSPMAGSLFKALMESGVAILPIRAPDDRNHDIQYIAR
 ICGNVSVDVALLOCTIRAKSSBELDIINKTKSFRVYVDGFFPFPBBDILTEKFPNS
 IPSVIGVNNHRCGLFLPKKEPPELIGSNKSLALHLIRVLHIENQYILVADQYFN
 KHSPEVIEKDFLDGDFVFPVPGVTVARHYHDAGAPVYFEPQHPBCLKQTRPAV
 KADHDEIRFVFGAFLKGNLYMFEAGATEBEKTLRSRKMRYANARAGCDPGEGLPL
 MPAYSQSEYQYKLDLNTISVGQKLKQEVPEFMSDTLPILMSMTAPRPPVPLSLSVL
 LPFLFSSAP"

ORIGIN

Query Match 68.6%; Score 1198.4; DB 4; Length 1728;

Best Local Similarity 82.7%; Pred. No. 0;
 Matches 1370; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 91 GGGCTTCTGCTGGAAGGCGACAGAGAACACACAGCTGGATGATTCAGGCGCAAGCAA 150
 DB 73 GGTCCAGCTGCGTAGAGCACCGGTGAGAGACACAGGCTGGATGGGTCCGGGCGAAGCAA 132
 QY 151 GTCACTGTGCTGGGAAGCCGTGTGCTGTGAACGTGTCTCTGGAAGTCCCTTGTCTGCT 210

Db 133 GCACTGTGTGGGAGACCATGCTGTGAAGTTCCTCGGAGATCCCTTGTGCG 192
 Qy 211 CCCCCGCTGGGATCCCTGCGATTTACGAACCCGAGCTGCATCGCCCTGGATACCTTG 270
 Db 193 CCCCCCTTAAGAACCCCTGCGATTTAAGAACCCGAAGCTGCATTTGCTGGAAATGACTCA 252
 Qy 271 CGAAGAGCACCTCTCACTCAATTTTGGCTCCAGAACCTCAGAGTGGCTCTTATGAT 330
 Db 253 CGAATGTGCAGTCTTATCTTAATTTGTGCTCCAGAACCTCAGTGTGCTCTCAAT 312
 Qy 331 CAACACATGCTCAAGTGCATTTACCGAATTCGAGTGTCAAGAACCTGCTCACTG 390
 Db 313 CAACACTTCTCAAGTGCATTTACCCCAATTTAAGATGTCCGAAGCTGCTGTACTT 372
 Qy 391 AACATCATGGGCTGCGCCAGCCGATACAGGCTCCAGCTCCCGCTTGTGTGTGTT 450
 Db 373 AACATCTACGACCAAGCCAGCCGAACACCGGCTCCCAAGCTCCCGTAATGTGTGTT 432
 Qy 451 CCAAGAGTGCCTTCAAGACTGTGCTCAGCTCCATCTTTGATGAGGTCCGCTGCGTGC 510
 Db 433 CCTGAGAGTGCCTTTGAGACTGTGCTCAGCTCCATCTTGAATGGTCCGCTGCTGCC 492
 Qy 511 TATGAGAGCTGTGCTGT 570
 Db 493 TATGAGATGT 552
 Qy 571 ACATGGAGTCAAGATCTCCGCGGAACTGGGCTTCAAGAACAGGAGTGTGTGTGTGT 630
 Db 553 ACAGGGGACAGACCGCTCCGCGGAACTGGGCTTCTGACCAAGTGTGTGTGTGTGTGT 612
 Qy 631 TGGGTCCAGAAACATCGAGTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 690
 Db 613 TGGGTCCAGAAACATCGAGTTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 672
 Qy 691 GAGTCCGCGGAGCCATAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 750
 Db 673 GAGTCCGCGGAGCCATAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 732
 Qy 751 TTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGAGAGCCCATGAT 810
 Db 733 TTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTCTGTGGGCCCCCAT 792
 Qy 811 TATGAGAGAGTGAAGACCTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 870
 Db 793 GATGAGAGAGTGAAGATTTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 852
 Qy 871 TCTGAGGCTCTGTGAGGTGCTGAGACAAACCTCCAAAGAGCTGTGACCTGTGAGC 930
 Db 853 TCCGTAGCTGTGTGCAAGTGTGCTGAGGCAAAATCTTCGAGAGTGTGTGACATCAAC 912
 Qy 931 CAGAAAACAAAGTCTTTCATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 990
 Db 913 AAGAAAACAAAGTCTTTCATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 972
 Qy 991 GATCATTTGTCTCAGAAAGCATTTTAAAGCAATCTTCATCATGTGAGTCAATTAACAC 1050
 Db 973 GACCTATTGACTGAGAAAACATTTAATTCATTTCTGTCAATGAGTCAATTAACAC 1032
 Qy 1051 GAGTGTGCTTCTGTGCTGTGCTGTATAGAGAGGCTCTGAGATCTCAGTGTCTCCAAAG 1110
 Db 1033 GAGTGTGCTTCTGTGCTGTGCTGTATAGAGAGTTCGAGAGATCTTGTGAGGCTCCAAAG 1092
 Qy 1111 TCCCTTGGCCCTCACTGTGATCAAAACATCTGTGACATCCCGCTCAGTATTGTGACCTT 1170
 Db 1093 TCTCTGGCCCTGACCTGTGATCAACCGAGTCTGTGACATCCCAACAGTACTTATATCTT 1152
 Qy 1171 GTGCTTATGATATCTTTCATGACAGACATCTCTGACTGAAATCCGAGACAGTCTTCTG 1230
 Db 1153 GTGCTTATGATATCTTTCATGACAGACATCTCTGACTGAAATCCGAGATGAGTATGTTCTG 1212
 Qy 1231 GACTGTGCTGAGATGTGTTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1290

Db 1213 GACTTGTGAGAGTGTGTTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1272
 Qy 1291 GATGTGTGTGACCTGTGTCACTTTATGATTTTGGCACCGGCTTCAAGTGTGTGAAGAC 1350
 Db 1273 GATGTGTGTGACCTGTGTCACTTTATGATTTTGGCACCGGCTTCAAGTGTGTGAAGAC 1332
 Qy 1351 ACGAAGCCGCTTTTGTCAAGACCGACACGCTGTATGAAGTCCGCTTGTGTGTGTGTGT 1410
 Db 1333 ACGAAGCCGCTTTTGTCAAGACCGACACGCTGTATGAAGTCCGCTTGTGTGTGTGTGT 1392
 Qy 1411 GCTTCTCAAGAGGAGGACATTTGTATTTTCAAGAGCCACGAGAGAGAGATTTACTG 1470
 Db 1393 GCTTCTCAAGAGGAGGACATTTGTATTTTCAAGAGCCACAGAGAGAGAGATTTGCTG 1452
 Qy 1471 AGCCGAGATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1530
 Db 1453 AGCAGAGATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1512
 Qy 1531 GACTGTCTGTGTGACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1590
 Db 1513 GACTGTCTGTGTGACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1572
 Qy 1591 ATGAGCTCTGAGACAGACTCAAGAAACCGCGGTGATTTTGTGACACAGACCATCTCC 1650
 Db 1573 ATGAGCTCTGAGACAGAACTGAAAGAGAGAGAGTGTGATTTTGTGATTAATCTCTCC 1632
 Qy 1651 CTGATCTGTGTGCTTCCAGATGCTCCAGATGCTCTTCTTCTTAACTTCTCTCT 1710
 Db 1633 CTGATTAATGTCATGTCACAGACACCCCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1692
 Qy 1711 CTCTCCAGGCTTCTTCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1746
 Db 1693 GTGCTCTGTGCTTCTTATTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1728

Search completed: June 14, 2005, 22:52:51
 Job time : 7589.61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 5793.46 Seconds
(without alignment)
11471.585 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 1746

Sequence: 1 atgcacacagagactacttc.....ttttctttgtgcctctga 1746

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: gb_ests1:*
2: gb_ests2:*
3: gb_hnc:*
4: gb_ests3:*
5: gb_ests4:*
6: gb_ests5:*
7: gb_ests6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	27.4	742	4	BI822069 603039938
2	418.6	24.0	1680	9	AY414461 Mus muscu
3	414.6	23.7	576	4	BI824830 60303738
4	406.6	23.3	1680	9	AY414459 Homo sapi
5	397.8	22.8	2661	3	AK033563 Mus muscu
6	396.2	22.7	2687	3	AK077248 Mus muscu
7	387.6	22.2	4927	3	AK040349 Mus muscu
8	383.6	22.0	1958	3	AK078953 Mus muscu
9	355.2	20.3	1944	3	AK078879 Mus muscu
10	345.8	19.8	2888	3	HSB803013 Homo sapi
11	344.2	19.7	3909	3	HSB806270 Homo sapi
12	338	19.4	1918	7	CP111083 Shultzom
13	331.4	19.0	2038	3	BC019926 Mus muscu
14	298.4	17.1	1826	3	CR618303 Ttll-1eng
15	293.6	16.8	1859	3	CR641368 Tetradon
16	292	16.7	1817	3	CR634885 Tetradon
17	289	16.6	1610	3	AK050444 Mus muscu
18	286	16.4	2356	3	AK037191 Mus muscu
19	285	16.3	491	5	BX280446 BX280446
20	284.2	16.3	1849	3	CR677187 Tetradon
21	284	16.3	539	5	BU585267 71701765V
22	284	16.3	593	1	AI808985 w67d05.x
23	282.2	16.2	805	5	BP442124 BP442124
24	282.2	16.2	808	5	BP446194 BP446194

25	282.2	16.2	840	5	BP445396	BP445396
26	282.2	16.2	841	5	BP447764	BP447764
27	282.2	16.2	860	5	BP441209	BP441209
28	282	16.2	904	5	BO900930	AGENCOURT
29	281.8	16.1	798	5	BP444116	BP444116
30	281.8	16.1	808	5	BP443157	BP443157
31	281.8	16.1	877	5	BP441564	BP441564
32	281.8	16.1	877	5	BP442855	BP442855
33	281.6	16.1	1823	3	CR640788	CR640788
34	280.8	16.1	1853	3	CR723405	Tetradon
35	280.2	16.0	840	5	BP442111	BP442111
36	279.8	16.0	783	5	BP445505	BP445505
37	279.6	16.0	792	5	BP443197	BP443197
38	276	15.8	645	5	BU585270	BU585270
39	275.4	15.8	905	7	CO583389	ILUWIGEN
40	274.8	15.7	989	3	AK007235	Mus muscu
41	274	15.7	1857	3	CR637401	Tetradon
42	273.6	15.7	1117	7	CN061747	Ag2_P32_1
43	272.4	15.6	763	7	CR765061	DKF2P470L
44	272	15.6	770	5	BP444251	BP444251
45	271.6	15.6	906	5	BU849807	AGENCOURT

ALIGNMENTS

RESULT 1
BI822069
LOCUS 603039938P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181056 5',
DEFINITION mRNA sequence.
ACCESSION BI822069
VERSION BI822069.1 GI:15933619
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 742)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1A011451 row: k column: 09
High quality sequence stop: 740.
Location/Qualifiers
1..742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5181056"
/lab_host="DH10B"
/clone_1lb="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match	27.4%	Score 479;	DB 4;	Length 742;
Best Local Similarity	100.0%;	Pred. No. 1.4e-125;		
Matches 479;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.

Qy	91	GGGCGTCTTCTGCTAAGGGGCGCAGAGGAAACACAGAGCTGGATGATTCAGGGCGAACCA	150
Db	103	GGGCGTCTTCTGCTAAGGGGCGCAGAGGAAACACAGAGCTGGATGATTCAGGGCGAACCA	162
Qy	151	GTCACTGTGCTGGGAAGCCCTGTGCTGTGMACGTGTCTCGAGTGCCCTTTGGTGAGT	210
Db	163	GTCACTGTGCTGGGAAGCCCTGTGCTGTGMAAGTGTCTCGAGTGCCCTTTGGTGAGT	222
Qy	211	CCCCGCTGGGATCCCTGCGATTTAAGAACCCGCACTGCACTGCGCTTGGGATTAATTG	270
Db	223	CCCCGCTGGGATCCCTGCGATTTAAGAACCCGCACTGCACTGCGCTTGGGATTAATTG	282
Qy	271	CGAGAAGCCACTCTCAACCTAAATTGTGCTCAGAACTCAGAGTGAGTGCTCTTAAGT	330
Db	283	CGAGAAGCCACTCTCAACCTAAATTGTGCTCAGAACTCAGAGTGAGTGCTCTTAAGT	342
Qy	331	CAACACATGCTCAAGGTGCTATTACCGGAATTCGAGTGTCAGAAAGCTGCTCTTAAGT	390
Db	343	CAACACATGCTCAAGGTGCTATTACCGGAATTCGAGTGTCAGAAAGCTGCTCTTAAGT	402
Qy	391	AACATCTATGCGCTGCGCCAGCGCGATTCAGGCTCCAGAGTCCCGCTCTTGTGTGCTTC	450
Db	403	AACATCTATGCGCTGCGCCAGCGCGATTCAGGCTCCAGAGTCCCGCTCTTGTGTGCTTC	462
Qy	451	CCAGAGGTGCTTCAAGACTGGCTCAGGCTTCATCTTTGATGGGTCCGCGCTGGCTGCC	510
Db	463	CCAGAGGTGCTTCAAGACTGGCTCAGGCTTCATCTTTGATGGGTCCGCGCTGGCTGCC	522
Qy	511	TATGAGAGCTGCTGCTGTGTGTGTGCTCAAGTACCGGCTAAGGAATTTGTGTTTCTTAC	569
Db	523	TATGAGAGCTGCTGCTGTGTGTGTGCTCAAGTACCGGCTAAGGAATTTGTGTTTCTTAC	581

RESULT 2	LOCUS	DEFINITION
AY414461	AY414461	1680 bp DNA linear GSS 17-DEC-2003 Mus musculus Ccsg gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

VERSION	KEYWORDS
AY144461.1	GI:39770423
SOURCE	ORGANISM
Mus musculus Mus musculus (house mouse)	
REFERENCE	AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus 1 (bases 1 to 1680) Clark,A.G., G lancowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeallo,D.R., Lu,F., Murphy,B., Periers,S., Wang,G., Zheng,X.H., White,T.D., Smithsky,J.C., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios <i>Journal of Molecular Evolution</i> 50(2): 1-13 (1999)	
TITLE	JOURNAL
gene trios	Science 302 (5652), 1960-1963 (2003)

gene <1..>1680
/gene="CS2"
/locus_tag="HCM5231"

Query Match	24.0%;	Score 418.6;	DB 9;	Length 1680;
Best Local Similarity	56.9%;	Pred. No. 3.8e-108;		
Matches 885;	Conservative 0;	Mismatches 634;	Indels 36;	Gaps 5

OY	115	AGGAACACAGGCTGGGATGGATTTCAGGGGCAAGAACATCTGTGTGGGAAGCCCTGTG	174
Db	100	AGGACCAACACAGAGCCAGATACCTGGGAGCCTTATCATATGAAGACCTGTATGTG	159
OY	175	CTGTGAACGTGTTCCTCGGAGTCCCTTTGTGTGCTCCCGCTGGATCCCTGGATTT	234
Db	160	GGTGTCCACTTCTCTGGGAAATCCCTTTGGCAAGCCCTGTGAGAACATGCGCTTT	219
OY	235	ACGAACCCGAGCTGCATTCGCCCTGGGATAACTTGGAGAAAGCACTCTCACTTAAT	294
Db	220	GCACCCCTTGAGCACCTGAACCATGGGGTGTGTGAAGATGGGACATCAATCCGCC	279
OY	295	TTTGTGCTTCAG---AACTCAGATGGCTGCTTTAGATCAACACATGCTCAAGGTGAT	351
Db	280	ATGTGTCTCAGGACATCACTGCAATGAATATGAGGCTTTTAAGCTTTAAGCTGAACC	339
OY	352	TACCGGAATTTGGAGATGTCAAGAAAGTCCCTCTACTGAAACATATATGCGCTGCCAC	411
Db	340	TTGCTTCCTCCCTTAATGTCTGAGAACCTGCTTAATCTTAAACATCTATGACACGATCAT	399
OY	412	GCCGATACAGGCTCCAAAGTCCCGCTCTGGTGTGTGCCAGAGGAGCCTTCAAGACT	471
Db	400	GCCCATGAAGGTTTCCAACTGCTGTGATGTGTGTGATTCACAGGTGTGAATCTGTGCA	459
OY	472	GGCTAAGCTCCATCTTTGATGTGGTTCGGCCCTGCTGTCTATGAGACGTGTGGTTTG	531
Db	460	GGGATGTCTTCATATATATGATGATCCCTGTGACACCACTGAGATGTGTGTGTGTC	519
OY	532	GTGCTCAGTACCGGCTAGAAATATTTGGTTCTTCAACAATGGGATTCAGCAATGCTCCG	591
Db	520	ACTATTCAGTACCGCTGGGTGTCTGTGGTTTCTTCAAGACATGGGACCAAGCATGTCCAGA	579
OY	592	GGGAACTGGGCTTTCAGGAACCAAGTGGCTGTCTGTCTGTGGTCCAGAAACATTCGAG	651
Db	580	GGCAATTTGGGATATCTGGAGCAAGTGGCTGTCTGCGTGTGGTCCAGGAATATGCGC	639
OY	652	TTCTTCGGTGGGAGCCCGAGCTGTGTACATCTTTGGGCAAGTCCCGGGAGGCAATTAAGT	711
Db	640	CACTTTGGTGGCAACCTGACCCGGTTTACAATTTTGGCGAGTCTCAAGTGGCACAAAGT	699
OY	712	GTTTCTAATCTTAACTGTCTTCCCATGGCCAAAGGCTTATTCACAAAGCATCATGGAG	771
Db	700	GTGTCTTCACTTGTGTGTGTCCCAATGTCCAAAGGGCTTTCATGTGTGCATCATAGAG	759
OY	772	AGTGGGGGTGCATCATCCCTTAACCTGGAGAGCCCATGATATATGAGAAGTGAAGACTTG	831
Db	760	AGTGGGGGTGCCTGTCTGCTTCACTTATTTCCAAACCTCTGAGA-----TGTGTCC	813
OY	832	CAGGTGTTCACATTTCTGTGTGAACAATGCTGTAGACTTGAAGGCTCTGTGTAGGTGC	891
Db	814	ACTACGGTGTGCAAAGCTCTCTGTGATGTAGGCGGTGAGCTCAGAGGCTCTGTGTGCTGT	873
OY	892	CTGAGGACAAACCCCTTCCAAAGAGCTGTGACCTCTAGCCAGAAACAAAGTCTTTCATCT	951
Db	874	CTGAGAGGCAAGATGAAGCAGAGATTTGGCTATTTAACAAACCTTCAGATGTATCCCT	933
OY	952	CGAGTGTATGATGTCTTCTTCTTAAATGAGCCTGTGATCTTATTTGTCTCAGAAACCA	1011
Db	934	GGTGTGTGTGATGGAAGATTTCTTAACCAAGCATCTCTCAGAGACTCTGTGCTGTGTAT	993
OY	1012	TTTAAAGCAATTCCTTTCATCATCGGAGTCAATAAACCAAGTGTGTGCTCTGTGTGCT	1071
Db	994	TTTACCCCTGTCCCGAGATCATTTGGTTTTCACAAATGATGTGATGTGTGATGCTGCC	1053

QY 1072 -----ATGAGAGAGGCTCTGAGATCTCAGTGGCTCCAAAGTCCCTTGGC 1119
DB 1054 AAAAGATGAGATGATTTGATCTCAAGAGGGGATGGAAGCAATGGTGCAGCTATC 1113
QY 1120 CTCATCTGATACAAACATCTCGACATCCGCTCAGTATTGGACCTTGGCTAAT 1179
DB 1114 CTGCAGAGAGGGCAACCCAGATGATGCTGCTCCCTGATGATGACCTGCTAATGGA 1173
QY 1160 GAATCTTCATGACAGACATCCCTGATGAAATCCGAGAGAGTCTTGGATCTTGGCT 1239
DB 1174 GATGACATGGGGGACCTGAGATGCCAGACCTCCAAATACAGTTCAAGAGATGATG 1233
QY 1240 GAGATGATGTTCTTGTGGTCCCTGACATGATCAGAGTCAATACAGAGATGCTGT 1299
DB 1234 GGAATTTTCAATGTTGATGATCTCTGCACTCCAGATGACATTTTCAAG---TTCCAT 1290
QY 1300 GCACCTGCTACTTCTATGATGTTTGGGACCGGCTCAGTGTGGAAGACAGAGCCG 1359
DB 1291 GCCCTGCTACTTCTATGAAATTCACATCAATCAGCTTCTCAAGATGTCAGGCCA 1350
QY 1360 GCTTTTTCAGAGCCGACCGCTGATGAGTCCGCTTGTGTGGTGGTCCCTTCTG 1419
DB 1351 CCCACGGAACGCTGACCATGAGATGAGTCTTGTGTGGTCTTCTTCTG 1410
QY 1420 AAGGGGCACTTGTATGTTTGGAGAGCCAGAGAGAGATTAAGTCTGAGCCGAG 1479
DB 1411 GGCATGAACCTTGACTT-----CACTGAGAGAGAGAGAGCTGTGAGAGAG 1458
QY 1480 ATGATGAATATCTGGGCTTACTTGTCTCGAACCGGAAATCTAATGGAGACGCTGTCT 1539
DB 1459 ATGATGAATATCTGGGCTTACTTGTCTCGAACCGGAAATCTAATGGAGAGCTGTCT 1518
QY 1540 CTGTGGCCAGCTTAATATCTGATGAGAGTACCTTCAGCTGGAATTAAGATGAGCTC 1599
DB 1519 TACTGGCCCTGATGATGACCATGATGAGAGTACCTTCAGCTGGAATTAAGATGAGCTC 1578
QY 1600 GACACGAGACTCAAGAACCGCGGGTGAATTTTGGACGACGACCTCCCTG 1654
DB 1579 GGTGACGCCCTGAGAGCCAGAGGCTGCACTTCTGACCAAGATCTGTGCCAG 1653

RESULT 3
BI824830 576 bp mRNA linear EST 04-OCT-2001
LOCUS 6030373881 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174937 5',
DEFINITION mRNA sequence.
ACCESSION BI824830
VERSION BI824830.1 GI:15936380
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 576)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1435 row: 1 column: 10
High quality sequence start: 6
High quality sequence stop: 565.
Location/Qualifiers

FEATURES
SOURCE
1..576
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:5174937"
/lab_host="BDH10B"
/clone_id="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT; Site: 1; NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC library."

Query Match 23.7%; Score 414.6; DB 4; Length 576;
Best Local Similarity 96.4%; Pred. No. 3.8e-107;
Matches 456; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 91 GGGCTTCTGCTGAGAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAGCA 150
DB 104 GGGCTTCTGCTGAGAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAGCA 163
QY 151 GTCACTGTGTGGAGAGCCCTGTGCTGTAAGCTGTTCTCTGAGTCCCTT-TGCTGC 209
DB 164 GTCACTGTGTGGAGAGCCCTGTGCTGTAAGCTGTTCTCTGAGTCCCTTATGCTGC 223
QY 210 TCCCTCGCTGGAGATCCCTGCAATTTAGAACCCGACGCTGATGCTCCCTGGATTA 269
DB 224 TCCCTCGCTGGAGATCCCTGCAATTTAGAACCCGACGCTGATGCTCCCTGGATTA 283
QY 270 GGGAGAACCCACCTCTCAATCTTATTTGTGCTCCAGAACTCAAGTGGCTCTTGA 329
DB 284 GGGAGAACCCACCTCTCAATCTTATTTGTGCTCCAGAACTCAAGTGGCTCTTGA 343
QY 330 TCAACATGCTCAAGAGTCAATTAACCGAAATTCGG-AGTGCAGAGAGTCCCTCA 388
DB 344 TCAACATGCTCAAGAGTCAATTAACCGAAATTCGGAGTGCAGAGAGTCCCTCA 403
QY 389 TGAACATGCTCAAGAGTCCCAAGCCGATACAGGCTCCAGCTCCCTTGTGTGT 448
DB 404 TGAACATGCTCAAGAGTCCCAAGCCGATACAGGCTCCAGGCTCCCTTGTGTGT 463
QY 449 TCCCGAGAGTCCCTTCAAGAGTGTCTCAAGCTCTCAATCTTGTATGGTCCGCTG 508
DB 464 TCCCGAGAGTCCCTTCAAGAGTGTCTCAAGCTCTCAATCTTGTATGGTCCGCTG 522
QY 509 CCTATGAGAGCGTGTGTTGTGTCGTCAGTACCGGCTAGGAATTTTGT 561
DB 523 CTATGACGAGCGTGTGTTGTGTCGTCAGTACCGGCTAGGAATTTTGT 575

RESULT 4
AY144459 1680 bp DNA linear GSS 17-DEC-2003
LOCUS AY144459
DEFINITION Homo sapiens CES2 gene, VIRTUAL TRANSCRIPT, partial sequence.
ACCESSION AY144459
VERSION AY144459.1 GI:39770421
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1680)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tenenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.D., Snitsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302
 REFERENCE 2 (bases 1 to 1680)
 AUTHORS Clark, A.G., Ghanowski, S., Nielson, R., Thomas, P., Kejaritwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
 Ferris, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1. 1680
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..21680
 /gene="CBS2"
 /locus_tag="HCM5231"
 ORIGIN
 Query Match 23.3%; Score 406.6; DB 9; Length 1680;
 Best Local Similarity 56.5%; Pred. No. 1.1e-104; Mismatches 639; Indels 36; Gaps 5;
 Matches 876; Conservative 0;

QY 833 AGTGGTTCACATTTCTGTGTAA CAA TGCCTCAGACTGTAGGCCCTGTGAGTCC 892
 DB 815 CGGTGGGCGCAACCTGTCTGCTGTGACCAAGTTGACTGTAGGCGCTGGTGGCTCC 874
 QY 899 TGAAGACAAACCTCCAGAGAGCTGTGACCTCAGCAGAAAAAAGTCTTCACTC 952
 DB 875 TCGGGGCAAGATTAAGAGAGATTTGCAATTMAACCTTTCAGATGATCCCCG 934
 QY 953 GAGTGGTATGTGCTCTTCTTCTTAATGAGCCCTGATCTATTTGTCTAGAAAGCAT 1012
 DB 935 GAGTGGTATGTGCTCTTCTTCTTAATGAGCCCTGATCTATTTGTCTAGAAAGCAT 994
 QY 1013 TTAAGCAATTCCTTCCATCATGAGTCAATTAACCAAGATGTGCTTCTGTGCTTA 1072
 DB 995 TTAAGCAATTCCTTCCATCATGAGTCAATTAACCAAGATGTGCTTCTGTGCTTA 1054
 QY 1073 TGA-----AGGAGCTCTGATGATCCCTCAGTGGCTCCACAAAGTCCCTTGGCC 1120
 DB 1055 AGTCAATGAGATCTATGATATCCAGAAAGAAATGACAGAGAGCCCTCCAGGCTGCTC 1114
 QY 1121 TCCATCTGATCAAAACATCTCTGACATCCCGCTCAGTATTTGCACTTGTGCTAATG 1180
 DB 1115 TGCAAAATGTTAAGCTGTGATGTTGCTCTTACATTTGTGACCTGTGAGGAGG 1174
 QY 1181 AATACTTCATGACAGACATCTCTGATGAAATCCGAGACATCTTGTGACTTGTG 1240
 DB 1175 AGTCAATGAGAGATGAGAGAGATCCCAAGACCTTCAAGGCGCAAGTTCAGAGATGATG 1234
 QY 1241 GAGATGTTCTTGTGTGCTCTGCTGACATGATCAAGCTGATATCAAGATGCTGTG 1300
 DB 1235 CGGACTCAATGTTGTGATCCCTGCACTCCAGTACAGACATTTTCACTG--TTCCGGG 1291
 QY 1301 CACTGTCTACTTATGATTTGCGACCGGCTCAGTCTTTGAAGACAGAGCCGG 1360
 DB 1292 CCCCTGTGATCTTCAAGATTCAGATCAGCTGCTGCTCAAGAACATCAGGCTAC 1351
 QY 1361 CTTTGTCAAGCGCAGACCTGATGATGAGTCCGTTTGTGTTCCGTTGCTTCTGTA 1420
 DB 1352 CGGACATGAAGGACAGACATGATGATGCTCTTGTGTTCAAGATTTCTTGGGG 1411
 QY 1421 AGGGGACATTTGATGTTGGAAGAGCCACGAGAGAGAGAAATTAAGAGCCGGAAGA 1480
 DB 1412 GCAACTACATTAATTT-----CACTAGGAAGAGAGACGATMAAGAGAGAGA 1459
 QY 1481 TGAATGAATCTGAGCTACTTGTGCTGAAACCGGAAATCTTAATGGAACGACTGTCT 1540
 DB 1460 TGAATGAATCTGAGCTACTTGTGCTGAAACCGGAAATCTTAATGGAACGACTGTCT 1519
 QY 1541 TGTGGCAGCTTATTAATCTGATGAGAGATCTCCAGCTGAGCTTGAACATGAGCTGG 1600
 DB 1520 ACTGGCGCTGTTCAGACAGAGAGAGACATCTGACGCTGAACCTACAGCTGCGGTG 1579
 QY 1601 GACAGAGATCAAAAGAACCGGGGTGATTTTGAACAGCAACATCCCC 1651
 DB 1580 GCCGGCTCTGAAGGCCCAAGGCTCAAGTCTGGAAGAAAGCGCTGCCCC 1630
 RESULT 5
 AK033563
 LOCUS AK033563
 DEFINITION Mus musculus adult male colon cDNA, RIKEN full-length enriched
 library, clone:9030624L02 product:similar to CARBOXYPESTERASE
 PRECURSOR (EC 3.1.1.1) (AII-ESTERASE) (B-ESTERASE) (MONOBUTYRASE)
 (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
 [Mesocricetus auratus], full insert sequence.
 AK033563
 VERSION AK033563.1 GI:26329244
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 6 (bases 1 to 2661)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, H., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ono, M., Ohsato, N., Okazaki, Y., Saito, R., Saiton, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
SOURCE 1. 2661
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:9030624L02"
/db_xref="taxon:10090"

/clone="9030624L02"
/sex="male"
/tissue_type="colon"
/tissue_lib="RIKEN full-length enriched mouse cDNA library"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
25. .1704
/note="unnamed protein product; putative
similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)
(ALI-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE
ESTERASE) (PROCAINE ESTERASE) (METHYLUTRASE)
[Mesocricetus auratus] (SPR|035533, evidence: PASTY,
83%ID, 100%length, match=1677)"
/codon_start=1
/protein_id="BAC28361.1"
/db_xref="gi:26329245"
/translation="MPYLKLGMLNANVACGVLLVLYHVGODSASPIRNTHTGVRGS
LVHVDIAHTFVIGIPAPCPVGPPLAPPAPEAPSPGVRDGTSHRMCLQNNLM
GSEDLKNNLLIPIGMSDELTYNLYVPAHAHESNLPLVWVHGLALVTGMASMD
GSMALATDVVVVAIQYRLGVLPSTDDHAKGWMGLDVAALRWQONIVHGVAN
PDRVTFESAGSVSSSHVSPMSQGFHAIKSVKIIIPGVYDGLPQHPQELMADHP
LSGCAAVNSFTLMCLRGKNEAMLAIKVKRIIPGVYDGLPQHPQELMADHP
VPSIIIGINDRYGMLPITMPDPAKIEITKTPAVYKSTALKMLPPEGCDLME
YMGDEDEFTIQAFREKMGDFVTPALQVAHFORSHPVYTFEFOHPSFDFRP
PYVADHGDDEIFLVFGYQGNIKLPTSEBQSLSRMKRYRANFARHGNPNSGLPYW
FVMDHDEQYQLDIDIPSVGRALKARRIQFTTKILPQKIOELKGSQERHKL"
2634. .2639
/note="putative"
2661
/note="putative"
ORIGIN
polyA_signal
polyA_site
Query Match 22.8% Score 397.8; DB 3; Length 2661;
Best Local Similarity 56.1%; Pred. No. 4,2e-102;
Matches 872; Conservative 0; Mismatches 647; Indels 36; Gaps 5;
115 AGGAACACGAGCTGGAGTGAATTCAGGCAAGCAAGTCACTGTCTGGAGACCCCTGTG 174
124 AGGAACACACACACAGGCAAGTGAAGGACGCTTGTCCAGTGAAGACACTGACATC 183
175 CCTGTGAACGTTCTCTGAGATCCCTCTTGTCTGCTCCCGCTGGATCCTGTGATTT 234
184 GCTGTTCATACCTCTCTGGAAATCCCTTTCGCAAGCCTCTGTGACCACTGGCGTTT 243
225 AGGAACCGGACGCTGCAATGCGCTGGAGTAACTTGCGAAGAACCACTCTCACTCAAT 294
244 GCACTCTTGAGGCCCCCTGAGCCATGAGTGGGTGAAGGATGGAACCTCAGATCCAAAC 303
184 GCTGTTCATACCTCTCTGGAAATCCCTTTCGCAAGCCTCTGTGACCACTGGCGTTT 243
225 AGGAACCGGACGCTGCAATGCGCTGGAGTAACTTGCGAAGAACCACTCTCACTCAAT 294
244 GCACTCTTGAGGCCCCCTGAGCCATGAGTGGGTGAAGGATGGAACCTCAGATCCAAAC 303
295 TTGTGCTTCAGAACTCAGAGTGGCTGCTTAA--GATCAACACATGCTCAAGGTGAT 351
304 ATGTGTCNAAAAATGAACAATTGATGGTTCAAGAGATCTGAAGATGATGAATTTGATC 363
352 TACCCGAAATTCGAGATGTCAGAGATGCTCTTACTGAAATTTATGCGCTGCCAC 411
364 CTGCTCCCATCTCATATCTGAGAGATGCTGTATCTCAACATCTATGTATCCAGCCAT 423
412 GCCGATACAGGCTCAACCTCCCGCTCTTGTGTGTTCCAGAGAGTGCCTTCAAGAT 471
424 GCCATGAGGGTTCTAACCTTACCTGTGATGATGATGATGATGATGATGATGATGAT 483
472 GGCTCAGCTCCATCTTGAATGGGTCCCGCTGCTGCTTATGAGAGAGTGTGTTGTG 531
484 GGCATGGCTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
532 GTGCTCAGTACCGGCTGAGAAATTTGTTTCTTACCAACATGAGATCAGATCTCG 591
544 GCTATCCAGTACCGGCTGAGTGTCTGAGATTTCTTACACATGAGAGACAGATCCCAA 603
592 GGGACATGGGCTTCAAGACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
604 GGCATGGGATACCTGAGACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
652 TTCTTCGTTGGGAGCCGAGCTGTGTGACATCTTTGGCGAGTCCGCGGAGCCATTAAT 711

Db 664 CACTTTGAGGCAACCTGACCGAGTCAACCATTTTGGAGAGTCAGCGGTGGACAAAGT 723
 Qy 712 GTTCTAGTCTTAACTGTCTCCATGGCCAAAGGCTTATTCACAAAGCATCATGAG 771
 Db 724 GTGTCTTCATATGTTGTGTCCCATCTCCAGGAGCTTTCATGAGTGCATCATGAG 783
 Qy 772 AGTGGGTGGCCATCATCCCTTACCTGGAGGCCATGATTATGAGAGATGAGACCTG 831
 Db 784 AGTGAAGTACTGTGTCTCCCTGATCTTATCTCCAGCTCCCTGAG------TGTTCAC 837
 Qy 832 CAGGTGTTGACATTTCTGTGTGTAAACATAGCTCAGACTCTTGAAGCCCTGCTGAGTGC 891
 Db 838 AGAATAGTGGCAATCATCTGTGTGTGAGCTGTGATGATCAAGACCCCTGATGTCTGC 897
 Qy 892 CTGAGGACAAACCTCCAAAGAGTGTGACCTCAGCCAGCAAAACAAAGTCTTCACT 951
 Db 898 CTAAAGGACGAAGATGAGAGAGATGTGTGTATTAATGAGTCTTCAAAATCATCTCT 957
 Qy 952 CGAGTGTGTAGTGTGTCTTCTTCTTAAAGACCTTAAGATCTATTTGTCTGAGAAAGCA 1011
 Db 958 GGTGTGTGATGAGAGATTCCTAACCAAGCATCTCAGAGAGTGTGCTCTAAGAT 1017
 Qy 1012 TTTAAAGCAATTCCTCCATCCGAGATCAATTAACAGAGTGTGCTTCTGCTGCT 1071
 Db 1018 TTTTCACTGTCTCCAGCATCATGTGTATCAAAATGATGAATATGTGTGAATCTTCTCT 1077
 Qy 1072 ATGAGAGAGCTCTCT-----GAGATCTCTAGTGTGTCCCAACAAAGTCCCTTGGCC 1119
 Db 1078 AGCATATCATGAGCCAGCTCAGAAATATAGAGAAATACAGAAAGACCTGTCACTGTT 1137
 Qy 1120 CTCATCTGTATCAAAACATCTCTGACATCCCGCTCAGATATTTGCACTTGTGCTAAT 1179
 Db 1138 CTAAAGAGCAGCCCTTAAATATGATGTCTCTCTGATGTGTGATGCTGTATGAG 1197
 Qy 1180 GAATATCTTCATGAGCAAGCACTCCCTGACTGAATCCGAGACAGTCTTGTGACTTGTGCTT 1239
 Db 1198 GAGTACATGGGGGAGACATGAGAGCCAGAGACTCTGCAAGCAAGTTCAGAGAGATGAG 1257
 Qy 1240 GAGATGTCTTTTGTGTGTCTCTGCACTGATCAGAGTCAATATCAGAGAGTGTGCT 1259
 Db 1258 GGGGACTTCATGTGTGTGTATCTCTGACTCACTCAAGATGAGCATTTTCA---ACGTTCCAT 1314
 Qy 1300 GCACCTGTCTACTTCTATGAGTTTGGGACCGGCTCAGTGTCTTGAAGACAGAGCGG 1359
 Db 1315 GCTCTGTCTACTTCTATGAGTTTCAACATGACAGCCAGCTTCTCAAGATTTTCAAGGCA 1374
 Qy 1360 GCTTTGTCAAGCGGACGAGCTGATGAGTCCGCTTGTGTGTGCTGTGCTTCTCTG 1419
 Db 1375 CCTATATGAGAGGCTGACATGATGATGAATTTTCTTGTGTGTGCTACCACTTGTGCT 1434
 Qy 1420 AAGGGGAGCATTTGTATTTGGAAGAGCCAGAGAGAGAGAGATTAATGAGCCGAGAG 1479
 Db 1435 AACATTAATCTCTCTTA-----CACTGAGGAGAGAGAGCAACTTAAGAGAGAG 1482
 Qy 1480 ATGATGAATATCTGGGCTACCTTTGTCTGGAACGGGGAATCTTATGGAAGAGACTGTCT 1539
 Db 1483 ATGATGAAGTACCGGGCCACTTTGTGACGAGCGGAAACCCCAAGAGAGAGTCACTCC 1542
 Qy 1540 CTGTGGCCAGCTTATTAATGATGAGAGAGTACCTCAGCTGGAATGAACATGAGCTC 1599
 Db 1543 TACTGGGCTGTGATGAGCATGATGAGAGATTAATCTGAGGCTGAGCATTCAGCTTCTGTG 1602
 Qy 1600 GGAAGAGAGCTCAAGAAACCGGGGTGATTTTGGACAGACCATCTCCCTGTA 1654
 Db 1603 GGTGAGCCCTGAAGGCGCAAGAGGCTCAGTTTGAACAAAGCTCTGCCCCAGA 1657

RESULT 6
 AK077248 2687 bp mRNA linear HTC 03-APR-2004
 LOCUS
 DEFINITION Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
 RIKEN full-length enriched library, clone:5031415B19
 product:similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)

COMMENT

(ALL-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
 (PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full
 insert sequence.
 AK077248
 AK077248.1 GI:26346111
 HTC; CAP trapper
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hazada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Maruyama, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6
 (bases 1 to 2687)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shitaki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 Direct SubMISSION
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genome Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail: genome-res@gscc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216]
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>

FEATURES

source

Location/Qualifiers

1..2687

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:5031415B19"

/db_xref="taxon:10090"

/clone="5031415B19"

/sex="female"

/tissue_type="ovary and uterus"

/clone_1ib="Riken full-length enriched mouse cDNA library"

/dev_stage="11 days pregnant adult"

53..1732

/note="unnamed protein product; putative

similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)

(AL1-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE

ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)

(Mesocricetus auratus) (SPTB|O35533, evidence: PASTY,

83%ID, 100%length, match=1677)"

/codon_start=1

/protein_id="BAC36707.1"

/db_xref="GI:26346112"

/translation="MPLVYLGLMNLNVAACVLLVLAHVGQDSASPIRHTHGOVRS

LHYVDTDAVHTFGLIPPAKPPVGLRPAPEAPBPMSGTDGSHPMCLQNDLM

GSIEDLKMNLILPISMSDELVLNIYVAHAHGSNLPVWVHIGALITVMSMYD

GSMLAEDLVVVAIOYRLGVLPFSTGQNAKNGVYLDOVAHLMVQNDVHSGN

PDRVITFGESAGTSVSHVAPMSQFLGHAHMGVAVLDLILSSSEMYHRIYAN

LSGCAVNSELMCLRGKNEBMLAINKFKIIPGVNDFEPLPKHPOBMSKDFHP

VPSILGINDEYGMILPTIMDPAOKIETRTTLPALVKSSTLKMMLPREGDILME

YMGEDPEPTLQAQPREMKGDPMFVITPAQVHFORSHAPVPIYERFORHPSFKP

PIVKAHDHDELPLVFGYQGNKILPYEEEDLSRIKMYANPAHAGNPSGAPYPM

PMMDHBOYLQIDQPSVGRALKARLQFWITLTPQKIQELNGSOERHKL"

2664..2669

/note="putative"

2687

/note="putative"

ORIGIN

Query Match 22.7%; Score 396.2; DB 3; Length 2687;
Best Local Similarity 56.0%; Pred. No. 1.2e-101;
Matches 811; Conservative 0; Mismatches 648; Indels 36; Gaps 5;

QY 115 AGGAACACCAAGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGAAACCCCTGTG 174
DB 152 AGGAACACACACAGGCGAGGTGAGAGCAGCTGTGCACGTGAAGACACTGACATC 211
QY 175 CCTGTGAAGCTTCTCGGAGTCCCTTTTGTCTGCTCCCGCTGGGATCCCTGGCATTT 234
DB 212 GCTGTCAACCTTCTCGGAAATCCCTTTGCCAAGCCCTCTGTAGAACACTGCGCTTT 271
QY 235 AGGAACCGGACGCTGCATCGCCCTGGGATTAATTGCGAAGAGCCACTCTAACCTAAT 294
DB 272 GCACTCTCTGAGGCCCTGAGCCATGAGAGTGGGCTGAGGATGGACCTTCAATCCAAC 331
QY 295 TTGTGCTTCAGAACTCAAGTGGCTGCTTAA--GATCAACACATGCTCAAGGTGAT 351
DB 332 ATGTGCTACAAATGACAAATTGAATGGGTCAAGGATCTGAAGATGAATGAATCTTGA 391
QY 352 TACCGGAATTTGGAGTGCAGAAAGTCTCTACTGAACATCTAATGCGCTGCGCCAC 411
DB 392 CTGCTCCCATCTCTAATGCTGAGAGCTGCTGTATCTCAACATCTATATACCAAGCCAT 451
QY 412 GCGATACAGGCTCCAAAGCTCCCGCTTGTGTGGTTCAGAGAGGTCTTCAAGACT 471
DB 452 GCGCATAGGGTTCTTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 511
QY 472 GCGTCAAGCTCAATCTTGAATGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
DB 512 GCGATAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 571

QY 532 GTGTCCAGTACCGGCTAGGAATATTGGTTTCTTACACATGGGATCAGATGCTCCG 591
DB 572 GCTATCCAGTACCGGCTAGGATTTCTGGATTTCTTACAGACTGAGACCAAGATGCAAA 631
QY 592 GGAACCTGGGCTTCAAGACCAAGTGGCTGTCTGTCTGGGTCCAGAAACATGAG 651
DB 632 GGAACCTGGGATATCTGGACCAAGTGGCTGTCTGTCTGGGTCCAGCAAAATTTGTC 691
QY 652 TTCTTGGGTGGGACCCAGCTCTGTGACATCTTTGGGATTCGCGGACCCATTAAT 711
DB 692 CACTTTGAGGCAACCTGACCGAATCACTTTTGGAGATCTGACAGTGGCACAAT 751
QY 712 GTTTCACTTAATCTGTCTCCATGCGCAAAAGCTTAATCCAAAGCCATCAATGAG 771
DB 752 GTGTCTCAATGTTGTGTCTCCCATGCTCCAGGACCTTCAATGATGATGATGATGAT 811
QY 772 AGTGGGTGGCCATCATCTTACCTGAGGCCATGATTAAGAAAGTGAAGACCTG 831
DB 812 AGTGGAGTAGTGTGCTCTCTGATCTTAATCTCAAGCTCCTGAGAG-----TGGTTCAC 865
QY 832 CAGTGTGTCACATTTCTGTGTAAATGAGTCAATGAGTCAAGTCAAGGCTGCTGAGTGC 891
DB 866 AGAATAGTGGCCAAATCTATCTGTGCTGCAAGTGTGAATCAAGACCTGATGTGCTGC 925
QY 892 CTGAGGACAAACCTCCCAAGAGCTGCTGACCTCAAGCAAAACAAAGCTTTCACT 951
DB 926 CTAAAGGCAAGAAATGAAGCAAGATGCTGCTATTAATTAAGTCTTCAAAATCATCCT 985
QY 952 CGAGTGTGATGATGCTTTCTTCTTAAGAGCTCTAATCTAATGCTCTGAAGAAAGA 1011
DB 986 GGTGTGTGATGAGAGATGCTTCAACCAAGCATCTCAGGAGCTATAGGCTCTAAGAT 1045
QY 1012 TTTAAAGCAATCTTCCATCATCGGAGTCAATTAACAGATGAGTCTTCTGCTGCT 1071
DB 1046 TTTCAACCTGCTCCAGCATATGATGATCAAAATGATGATGATGATGATGATGAT 1105
QY 1072 ATGAAGAGGCTCT-----GAGATCTCAAGTGTCTCAACAAAGTCCCTTGGC 1119
DB 1106 ACATATCAAGACCCAGCTCAAGAAATAGAGAAATACAGAAAGACCTGCAAGTGT 1165
QY 1120 CTCCATCTGATACAAACATCTGCAATCCGCTCAAGTATTTGCACTTGTGCTAAT 1179
DB 1166 CTAAAGAGACAGCCCTAAATATGATGCTGCTCTGATGATGATGATGATGATGATGAT 1225
QY 1180 GAATATCTCATGACAGACATCTCTGACTGAATCCAGACAGTCTTGTGACTTGTCT 1239
DB 1226 GAGTACATGAGGAGACATGAGGACCAAGACTCTGCAAGCAAGTTCAGAGATGAAG 1285
QY 1240 GGAATGTTCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1299
DB 1286 GGGAGCTTCAATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1342
QY 1300 GCACTGTCTACTTCTATGATGTTGCGACCGGCTCAAGTCTTGAAGACAGAGCGG 1359
DB 1343 GCTCTGTCTACTTCTATGATGTTGCAATGAGTCAAGTCAATTTTCA--AGTTCCTCAT 1402
QY 1360 GCTTTTGAAGACCGACAGCTGATGAAGTCCGCTTGTGTGTGTTGTGTTGTGTTGT 1419
DB 1403 CCTATGTAAGGCTGACATGATGATGAATTTCTTGTGTTGTGTTGTGTTGTGTTGT 1462
QY 1420 AAGGGGACATTTGATGTTGAAGAGCACGAGGAGGAGGAGTTTCTGAGCCGGAAG 1479
DB 1463 AACATTAATCTCTTCA-----CATTGAAGAGAGAGACATTAAGACAGAGAG 1510
QY 1480 ATGATGAATAATCTGGGCTAATCTTGTCTGAAACCGGAAATCTTAATGGAACAGCTGTCT 1539
DB 1511 AATAATGAAGTATCTGGGCAATTTGACAGGACCGGAAACCCCAACAGTGAAGGCTTACCC 1570
QY 1540 CTGTGGCAGCTTAATTAATGATGAGAGATCTTCAAGCTGAGCTTGAACATGAGCTTC 1599
DB 1571 TACTGGCTGTGATGAGACATGATGAGAGTACTGAGCTGAGCATCAAGCTTCTGTG 1630

QY 1600 GGACGAGACTCAAGAACCGGGGTGATTTTGGACGACCATCCCCCTGA 1654
Db 1631 GGTGAGCCCTGGAAGGCGCAGAGCTGCTGACACGACGATCTGCCCCAGA 1685

RESULT 7
AK040349

LOCUS
DEFINITION
AK040349 4927 bp mRNA linear HTC 03-Apr-2004
Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430088E12 product:similar to CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) (Mesocricetus auratus), full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK040349 GI:26087790
HTC; CAP trapper.
Mus musculus (house mouse)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
20530913
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
6 (bases 1 to 4927)
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirose, T., Hori, F., Imomori, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shikata, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, A., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE
JOURNAL
Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of

COMMENT
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216
URL: <http://genome.gsc.riken.jp/>,
URL: <http://fantom.gsc.riken.jp/>,
Location/Qualifiers
1..4927
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A430088E12"
/db_xref="taxon:10090"
/clone="A430088E12"
/issue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_abbrev="0 day neonate"
25..1711
/note="putative
similar to CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) (Mesocricetus auratus) (SPR103553, evidence: FASTY, 70.7%ID, 99.2%length, match=1672)"
4912..4917
/note="putative"
4927
polya_site
polya_site
misc_feature
misc_feature
25..1711
/note="putative
similar to CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) (Mesocricetus auratus) (SPR103553, evidence: FASTY, 70.7%ID, 99.2%length, match=1672)"
4912..4917
/note="putative"
4927
ORIGIN
Query Match 22.2%; Score 387.6; DB 3; Length 4927;
Best Local Similarity 56.1%; Pred. No. 4.3e-99;
Matches 890; Conservative 0; Mismatches 659; Indels 37; Gaps 7;
QY 90 TGGGCTTGTCTGTAAGGGCCACAGAGAACACACAGGCTGGATTCAGGCAAGCA 149
Db 105 TGAATCATGAGAGCCAGCCCATGAGAAACACATCTGACAGAGCTCGAGGCAAGTT 164
QY 150 AGTCACGTGCTGGAAGCCCTGCTGTAACGTGCTTCTGAGAGTCCCTTTGCTCG 209
Db 165 TGTTCACCTGACAGACATCAAACTGAGGCCATTAATCTTCTGGGATCCCTTCCAA 224
QY 210 TCCCGCTGGAGATCCCTGATTTAGAAC--CCGAGCTCATGCTCCCTGGATTA 268
Db 225 GCCTCTGTAGAGCACTGCGCTTTCACCTCTCTGAGGCCCTGAGCATGAGTGTG 284
QY 269 TCGAGAGCCACCTTCTTAACCTTAATTTGCTTCAGAACTCAGAGTGGCTCTT 328
Db 285 TGAAGATGGAACCTCAGACCGGCGCATGTGTCAAAATGATATGATGATTTAG 344
QY 329 A---TCAACATGCTCAAGGTGATTAACCGAATTTGGAAGTGCAGAAAGCTCC 385
Db 345 AGGGTGTGAAGATTAATAAAGATGATCTGCTCCCTTCTATGTCTGAGACTCC 404
QY 386 ACCGTAACATTAATGAGCTGCTGCGCAAGCGATAGAGCTCCAAAGCTCCCGCT 445
Db 405 ATCTCAACATTAACACACAGCCACGCGCAGAGAGGCTTAACTCTCTGATGTGT 464
QY 446 GGTTCACAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTTGATGAGTCCGCT 505
Db 465 GGCTCATGTGTGTGAGCTGTTGAGCAGCATGAGCTTCATGATGATGATCCGCT 524
QY 506 CTGCTATGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
Db 525 CAGCCATGAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 584

QY	566	TCACCAATGGGAATGAGATGCTCCGGGGAACTGGGCGCTTACAGGACCAAGGCGCTGCTC	625
Db	585	ACAGCACTGGAGATGACGAGCGCAAGGCACTGGGAAATCTTAGACCAAAACGCTGCC	644
QY	626	TGTCCTGGGTCAGAAAGACATCGAGTTCCTCGGTGGGAGCCCAAGCTCTGTGACATCT	685
Db	645	TACACTGGGGTCAGAGAAACATGGCCAACTTGGAGGCAACCTGATCTGTACACATTT	704
QY	686	TTGGCGAGTCCGGCGGAGCCCAATGATGTTTCTAGTCTATATCTGTCTCCCATGGCCAAAG	745
Db	705	TTGGCGAGTCAAGCAAGTGGCAACAGTGTGTCTTTTCATGTCTTGTCCTCCCGTGTCCAAAG	764
QY	746	GCTTATTTCCAAAGGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGGCC	805
Db	765	GACTCTTCCATAGAGCATCATGAGAGATGGGGGTGGCCCTGTGTCTACATTAATACCTG	824
QY	806	ATGATTTATGAGAAAGATGAGGACCTGCAAGGTGTTCACATTTCTGTGTAAACAATGGCT	865
Db	825	ACTCCCTCCGAAATGA-----TCTTTCACAAAGGTGGCAAAACCTATCTGTGTGAAACTT	878
QY	866	CAGACTCTGAGGGCCCTGCTGAGGTGGCTCGAGGACAAACCCCTCCAGGAGCTGTGACCC	925
Db	879	CAAACTCAGAAAGCCCTGTGTGGCTCTTCCAGGCAAGATGAGAGCAAAATTTCTGGCTA	938
QY	926	TCAGCCGAAAAACAAGCTCTTTCATCTCGAGTGTGTGATGTGTCTTCTTCTAATGAGC	985
Db	939	TGAGCAAGGCTTCAGGTTTATGCTGCTGTGTGTGATGGGAAATTCCTACCCAGACATC	998
QY	986	CTCTAGATCTATTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGGAGTCATA	1045
Db	999	CCAAAGATTTCTGGCTCTGTCTGATCTTTCACCTGTCTCCAGCATTTATGGTGTCAACA	1058
QY	1046	ACCAAGAGTGGCTTCTGTGCTGCTTATGAAGS-----AGGCTCTGAGATCC	1093
Db	1059	ATGATGATATGGCTGAGATTAATCCTTAAGATCTTCAATTTTCTCAGACAAATPAGGAAA	1118
QY	1094	TCAGTGGCTCAACAAGTCCCTTGGCCCTTCATCTGATCAAAACATCTGTGCATCCGCC	1153
Db	1119	TAAATAGAAATACCTGMAAGCTATTAAGAAATCAACAAGAGCAATGATGCTGCTT	1178
QY	1154	CTCAGTATTTTCACTTGTGGCTAATGAATCTTCCATGACAAAGCATCCCTGACTGAAA	1213
Db	1179	CTGAGTGTGTGACTGCTGTAATAGAAAGTACTTTAAGAGACATGAGAACCCCTGGACAC	1238
QY	1214	TCGAGACAGTCTTCTGGACTTCTTGGAGATGTGTTCTTTGTGGTCCCTGCATGTATCA	1273
Db	1239	TCCAAATTCAGATTGAGAGATGATAGGGGACTTCTTAATTAATCCCTGCATCCCAAG	1298
QY	1274	CAGCTCGATATCACAGAGATGCTGTGACCTGTCTTAATCTTATGATGTTTGGGACCGGCG	1333
Db	1299	TAGACAGTTTCAAGG--TTCCCATGTCTCTGTACTTCTTATGATGTTCAAGCAATAGGT	1355
QY	1334	CTCAGTCTTTGAAGACAGAAAGCCGCTTTTGTCAAGAGCCGACACACGCTGATGAATGCC	1393
Db	1356	CCAAGTTTACTCAAGTATTTCAAGGCCATGTGCAATGTAAAGCTGACCAACGCTATAGCTTT	1415
QY	1394	GCTTTGTGTTCGGTGTGCTCTTCTGAAAGGGGACATTTGTTATGTTCGAAGAGCCACGG	1453
Db	1416	ACCTTATCTT---TGGATCTTCTTCTTGGGGGCTCAAAATTTTAGCTT-----CACTG	1463
QY	1454	AGAGAGGAATTTTCTGAGCCGGAAAGATGATGAATACTGGGCTTACCTTGTCTGGAACGG	1513
Db	1464	CGAGAGAGAACTGTGAGCAGAAAGATGATGAATACTGGGCCCACTTTGACAGATATG	1523
QY	1514	GGAATCTTAATGGGAACGACTGTCTCTGTGGCCAGCTTATATTTGACTGAGCAGTATCC	1573
Db	1524	GGAACCCCAACAGTGAAGGATCTAACCTTACTGGCCAGCTCAAAACCAAGATATCTGTATCC	1583
QY	1574	TCCAGCTGGAATTTGAACATGAGCCCTCGGACAGAGCTCAAGAAACCGGGGCTGATTTT	1633
Db	1584	TGACGTGTGAATTCACCTTCTGTGTGGCCATGCTCTTAAGGCCAGAAAGCTGTCCATCTT	1643

Db	1644	GGACCAAGACTCTTCCCAAGATG	1669
RESULT 8			
LOCUS	AK078953		
DEFINITION	AK078953	1958 bp	RNA linear HTC 03-APR-2004
ACCESSION	AK078953		
VERSION	AK078953.1	GI:26347654	
KEYWORDS	HTC, CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS	1	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
AUTHORS	3	Normalizaton and subraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
TITLE	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20499374		
MEDLINE	11042159		
PUBMED	11042159		
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakauechi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
AUTHORS	4	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	
TITLE	Genome Res. 10 (11), 1757-1771 (2000)		
JOURNAL	20530913		
MEDLINE	11076681		
PUBMED	11076681		
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
AUTHORS	5	Functional annotation of a full-length mouse cDNA collection	
TITLE	5	Nature 409, 685-690 (2001)	
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			

COMMENT

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-research@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES

source

1. 1958
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9130231C15"
/db_xref="taxon:10090"
/clone="9130231C15"
/sex="male"
/isue_type="cecum"
/clone_id="Riken full-length enriched mouse cDNA library"
/dev_stage="adult"
66. 1742
/note="unnamed protein product; putative
similar to LIVER CARBOXYLSTERASE PRECURSOR (EC 3.1.1.1)
[Mesocricetus auratus] (SWISSPROT|Q6419, evidence: FASTA,
78.4kD, 100% length, match=1674)"
/codon_start=1
/protein_id="BAC37476.1"
/db_xref="GI:26347655"
/translation="MLARLPGMLCVACGLLLQLQHYHSDASPIRINRTHGOVRS
FVHVDTSGVAFGLIPKAPVGLLPEAPDEPSPGVDGTSOPALCPDIPNM
LEDAKEMNLLPPIGMSDECLYNTYTPHQAQESNLPMWIIHGGLVYSASANDY
SLAARERIVIVAIOYRLGVGFSTGQOHARGMNGYIDVVALRWKKNIAFRGNR
DRVITPGVASGTSVSHLSHSPKGLFHGALMOSGVALLPOLISDSEVYKTVANL
SGCEATDSBALHICLRASKDEILAIINOVFKMIPAVVGEBLPKIPOLILSMFHPV
PSIIIGVNTDECGVMFPMGLDHIKNTIRETLPAVLKNTAARMMLPESGHLVEXY
MGDTDEPTLQAQFREMGLDFWVIPAIVAHFORSOAPVYFYEFOHLSSTFKIVRS
HKADHDGDVAFVFGSYLMDMLDLTEEBELIKRMVMKYNFANPANGNPSGILPSPM
VLDDHBOYLOLDTPAVGVALKARLQFWITTLPLKQIDELGSDPKAEI."

polyA_signal

/note="putative"
1958
/note="putative"
/note="putative"

polyA_site

1958
/note="putative"

ORIGIN

Query Match 22.0%; Score 383.6; DB 3; Length 1958;
Best Local Similarity 55.1%; Pred. No. 4.5e-98;
Matches 855; Conservative 0; Mismatches 664; Indels 33; Gaps 4;

QY 115 AGGAACACCAAGCTGGATGATTCAAGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTG 174
DB 165 AGGAACACCCACAGGGGTCAAGTGAAGGACGCTTGTCCACGTGAAGACACTAAATCT 224
QY 175 CCGTGAACCGTGTCTCCGAGTCCCTTTGTCTCTCCCGCTGGGATCCCTGGATT 234
DB 225 GGTTCATGCTTCCCTGGGAATCCCTTGGCCAAAGCTCTGTAGGACTACGCGCTT 284
QY 235 ACGAACCCGAGCTGTGATCGCCTGGGATTAATTGGAGAAAGCCACTCTACCCCTAAT 294
DB 285 GCAACCTCTTAGAGACCTTAGACCATGAGTGTGAGAGATGGACCTCAAGCGGCC 344
QY 295 TTGTGCTTCAAACTCAGAGTGGCTCTTATGATCAACAATGCTCAAGTGCATTAC 354
DB 345 ATGTGTGAGCGCTGATATATGAATTTAGAGATCGAAGAAATGAATTCGATCCTG 404
QY 355 CCGAATTCGAGAGTGAAGAACTGCTTACTGAAACATCTATGCGCTGCCAGGCC 414
DB 405 CTCCTCATCTCTATGTCCAGAGACTGCTGTATCTCAACATCTAACACCAACCCAGCC 464

QY 415 GATACAGGCTCCAAAGTCCCGCTTTGTGTGTGTTCCAGAGAGTGCTTCAAGACTGCG 474
DB 465 CAGAGGGCTCTTAACCTGCTGTGATGATATGATATCCAGCGGTGGAGACTGGTTAGGA 524
QY 475 TCAGCTCCATTTTGAATGGGTCGCCCTGGTGTCTTATGAGAGACTGTGTGTGTC 534
DB 525 TACGCTTCATTAATGATGATATCTTAACCTGGAGGCACTGAGAAATGTATTTGTCC 584
QY 535 GTCCAGTACCGGCTAGGAATATTTGTCTTCAACACATGGAGATCAGATGCTCCGGG 594
DB 585 ATCCAGTATCGCTGGGGGCTTGTGCTTTCAGACATCGAGACACAGATCCAGAGCC 644
QY 595 AACTGGGCTTCAAGACCAAGGTGGCTGTCTGTCTGGTCCAGAGAACATCGAGTTC 654
DB 645 AACTGGGANTACCTGGACCAAGTGCTGCTTACCGCTGGGTGAGAAATATGCTTAC 704
QY 655 TTGGTGGGAGCCCGAGCTGTGACATCTTTGGCGAGTCCGGGGCCATAAGTGT 714
DB 705 TTGGAGGCAACCGAGACAGGCTCATATATTTGAGTCTGAGCAAGAGTAAAGTGTG 764
QY 715 TCTAGCTTATATCTGTCTCCATGAGCCCAAGGCTTATTCACAAAGCATCATGAGAGT 774
DB 765 TCTTCTCATATTTCTGTCCCAATGTCACAAAGACTCTTTCATGAGGCATTATGACAGT 824
QY 775 GGGGTGGCATCATCCCTTACCTGAGGCCCATATATGAGAAAGTGAAGACTCGGAG 834
DB 825 GGAATGGCCCTGTGCTGCTGACCTCATCTGTGACACTTCTGAG-----GTGCTTACAG 878
QY 835 GTGTTGCAATTTCTGTGTGTAACATGCGTCAGACTCTGAGGCTGTGAGGTGCTG 894
DB 879 ACAGTAGCAACCTATCTGTGTGTGAGGACAGACTCAGAGGCTGATTCACATGCTG 938
QY 895 AGGACAAACCCCTCCAGAGAGCTGCTTACCTCAGCCAGAAAACAAAGCTTTTCACTGA 954
DB 939 AGACAAAGATTAACCAAGAGATCTCGCTATTAAACCAAGCTTCAAGATGATCCCTGCT 998
QY 955 GTGTTGATGATGCTTCTTCTTCAATGAGCCTCTAGTATATGTCTCAAGAAAGATT 1014
DB 999 GTGTTGATGAGGAGATTCCTACCAAGACTCTCAGAGCTGTGACCTCTATGAGATTTT 1058
QY 1015 AAGCAATTCCTTCATCATGAGAGTCAATTAACCAAGAGTGTGCTTCTGCTGCTATG 1074
DB 1059 CACCTGTCCCGAGCATCATGTGTGTCAACAGATGATGTGCTGGGAGTCCCATG 1118
QY 1075 AAGAGGCTCTG-----AGATCTCAGTGGCTCCCAAGTCCCTTGGCCCTC 1122
DB 1119 TTCAATGGGCTTACCAATATCATTAAGAACATTAACAGAGAGCCCTGCAGCTGTCTG 1178
QY 1123 CATCTGATACAAACATCTGACATCCGCGCTCAGTATTTTGACCTTGTGCTAATGA 1182
DB 1179 AAGAACACAGACAGAGATGATGTGCTCCAGAAATGTATGTCACCTGTAATGAGAG 1238
QY 1183 TACTTCATGACAGACATCTCCTGACTGAATTCGAGACATCTTGTGACTTGTGA 1242
DB 1239 TATATGGGAGACCTGAGACCCAGACCTGCAAGCAAGTTCAGAGATGTGCTGGG 1298
QY 1243 GATGTCTTCTTGTGTGCTGCTGACATGATCAGAGCTCGATATCAAGAGATGCTGTGA 1302
DB 1299 GACTTATGTGTGTATCTCTGCTC---TCAAGTGGACATTTTCAGGCTTCCAGGCT 1355
QY 1303 CCTGTCTTACTTATGAGATTTGGGACCGGCTCAAGTCTTGTGAAGACAGAGCGGCT 1362
DB 1356 CCGTCTTACTTATGAGATTTCCAGCATGTGTCAAGCTTCATCAAGCATGTGAGGCATCA 1415
QY 1363 TTGTCAAAAGCCGACCAAGCTGATGAAGCGCTTTGTGTGAGTGTGCTTCTGAAG 1422
DB 1416 CATGTGAAGGACAGACATGATGATACGTTGCTTTGTGCTCTTATTTGTGGAC 1475
QY 1423 GGGGACATTTATGTTGGAAGAGCCACAGAGAGAGAAAGTTACTGAGCCGAGAAATG 1482
DB 1476 ATGAATTTGAGACT-----CACTGAGAGAGAGAGCTGCTTAAGAGATGATG 1523
QY 1483 ATGAATATCTGGGCTACTCTTGTCTGAACCGGGAATCTTAATGGAGAGACTGTCTGTG 1542

```

Db      1524 ATGAGTACTGGGCCAATTTCGACAGAAATGGGAACCCCAACAGTGAAGGTCTACCTCC 1583
Qy      1543 TGGCCAGCTTAAATTAATCTACTGACAGCTGCTCTCAGCTGAGATTGAACATGAGCCTCGGA 1602
Db      1584 TGGCTGTGTGTGGACCATGATGATGACGTACCTGACAGCTGAGAACCCAGCCTCTGTGGGT 1643
Qy      1603 CAGAGACTCAAGAACCCGGGGTGAATTTTGGACGAGCAACCATCCCTCTGA 1654
Db      1644 CGAGCCCTGAGGCCAGAAAGGCTGCACTTTCGACCAAGACTCTGCCCCAGA 1695

RESULT 9
LOCUS   AK078879                      1944 bp    mRNA    linear    HTC 03-APR-2004
DEFINITION Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030604P03 product:carboxylesterase 3, full insert sequence.
ACCESSION AK078879
VERSION   AK078879.1 GI:26347580
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE      High-efficiency full-length cDNA cloning
JOURNAL    Meth. Enzymol. 303, 19-44 (1999)
MEDLINE    99279253
PUBMED     10349636

REFERENCE
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
TITLE      Normalization and subtracction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL    Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE    20499374
PUBMED     11042159

REFERENCE
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuami,T., Teshiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Okazaki,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Yonezaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
TITLE      RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL    Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE    20530913
PUBMED     11076861

REFERENCE
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE      Functional annotation of a full-length mouse cDNA collection
JOURNAL    Nature 409, 685-690 (2001)
MEDLINE    11076861
PUBMED     11076861

REFERENCE
AUTHORS   The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE      Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL    Nature 420, 563-573 (2002)
MEDLINE    12000000
PUBMED     12000000

AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koyama,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,

```

```

Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M., and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT   cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://phantom.gsc.riken.jp/.
FEATURES
source     1..1944
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="PANTOM,DB:9030604P03"
            /db_xref="taxon:10090"
            /clone="9030604P03"
            /sex="male"
            /tissue_type="colon"
            /clone_1ib="RIKEN full-length enriched mouse cDNA library"
            /dev_stage="adult"
            64..1761
            /note="unnamed protein product; carboxylesterase 3 (MGD|MG1:2148202, GI|NM_053200, evidence: BLASTN, 99%, match=1939)
            putative"
            /codon_start=1
            /protein_id="BAC37439.1"
            /db_xref="GI:26347581"
            /translation="MGLPLILWLSLAACRWKPPSPPVNTVKGVLAKYNNLSEFT
            QVAVFLVGPFAKPLPLSLRFAFPQPAEWSVVKNTSTPPMCSQDAVGQVLSLFT
            NRKNTIPLQFSDCYLYNTYTPADLTPLRSLRWVWVHGGGLVVGASTYDGLASAH
            ENVVVVTIOYRLGIVFSTGDEHSGWGHLDQVLAALWVDNLNFGNGSGSVTF
            GSSAGGSVSVLPLAKNLFPHRAISSEGVLTALITDVKPIAGLVATSGCKTT
            TSAVWHCCRQKTDELETSKLNLPKILGLNGKESVPLPYDGVLPKAPBEI
            LAEKSPFVPIVGLNKQFEGWITITLNGYPLAEGLDKTANSLMKSYPFLKISEN
            MIPVAERKLAGTDLDLTKKKDLFDLMDVVGVSIVSRHDAGSTVWIEEYR
            SFSVAMRPKAVIGDHDEISFVFGSPFKDGASEETNLSDVWKFPMFANFANGNP
            GGLPHMPEYDQKEGYLKGISTQAORLKDKEVSFMELRAKESAPSHREHVEL"
            1922..1927
            /note="putative"
            1944
            /note="putative"
            /note="putative"
            /note="putative"
ORIGIN
Query Match      20.3%; Score 355.2; DB 3; Length 1944;
Best Local Similarity 55.1%; Pred. No. 6.3e-90;
Matches 870; Conservative 0; Mismatches 633; Indels 75; Gaps 6;
Qy      117 GAACACGAGCTGGATGATTCAAGGCAAGCATCTGCTGGGAAGCCTGTGCC 176
Db      141 GAACACTGTTAAAGCAAGTCTGGGGAAGTATGTCATTGGGAAGATTCAACAGCC 200
Qy      177 TGTGAACGTGTTCTCGAGAGTCCCTTGCTGCTCCCGCGCTGGAGTCCCTGGATTAC 236
Db      201 TGTGCGTGTGTTCTCGGAGTCCCTTGCCCAAGCCCCCTGTGGCTCTTGAAGATTGCG 260
Qy      237 GAACCGGAGCTCGATGATCGCTGGATTAAGGAGAGACACTCTACCTTAATT 296
Db      261 TCACACGAGCTCGAGAGCCTTGAGCTTGTGAAGAACACACACTCTACCGGCTTAT 320
Qy      297 GTGCTCCGAACTC-----AGATGGCTGCTTTAGATCAACATGCTCAAGGTGCA 350

```


Db 321 GTGCTCAGAGATGCTTGGAGGGCAGTGTCTCAGAGCTTTTCAACAAGGAAGA 380
 QY 351 TTACCCGAATTTGGAGGTCAAGACTGCTTACTGAAACATCTATGCGCTGCCA 410
 Db 381 GAACATTCCTTTACAGTTTCTGAAAGCTGCTTACTGAAATTTTACACTCTGTGTA 440
 QY 411 CGCGGATACAGGCTCCAAAGCTCCCGCTTGTGTGTGTCTCCAGAGAGTCTTCAAGAC 470
 Db 441 CTTGACAAAGAACAGACGACTACAGTGAATGTGTGATCATAGAGGTGATCTGTGTGT 500
 QY 471 TGGCTAGACCTCCATCTTTTATGTGGTCCGCTGTGCTGCTTATAGAGACTGTGTGTGT 530
 Db 501 GGGCGAGACATCCATCTATATGATGATGCTGCGCTCTCTGCGCCATGAAATTTGTGTGTGT 560
 QY 531 GGTGTCTCAGTACCGGCTAGAGAAATTTGTGTCTTCAACAATGGGATAGAGTCTCC 590
 Db 561 GACCATTTCAATACGCTTGGCATCTGGGATTTCTTCAAGCATGGGAGTAAACAAGTGTG 620
 QY 591 GGGGAACTGGGCTTCAAGAGCAGGTGGCTGTCTGTCTGTGGGTCCAGAAAGACATCGA 650
 Db 621 GGGAACTGGGGTCACTTGGACCAAGGTGGCTGCACTAAGCTGGGTCCAGGACAACTTGC 680
 QY 651 GTTCTTGGGAGGAGCCCAAGCTGTGTGACATCTTTGGGAGTCCGCGGAGCCATTAAG 710
 Db 681 CAACCTTTGGGAGGACCCAGCTGTGTGACATCTTTGGAGTCAAGAGTCAAGAGTTCAG 740
 QY 711 TGTCTTACTTATATCTGTCTCCATGGGCAAAAGGCTTATTCACAAGGCCATCATGGA 770
 Db 741 TGTCTGTCTTGTGTCTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 800
 QY 771 GAGTGGGGTGGCTATCATCTCTTCACTGGAAGGCCATGATTAAGAGAGTGAAGACCT 830
 Db 801 GAGTGGTGTGTCTCTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 860
 QY 831 GCAAGT 890
 Db 861 GGTTCCTACTCTTTCTGGGTGTAACTTCAACATCACTGTGTGTGTGTGTGTGTGTGTGT 920
 QY 891 CCTGAGACA-----AAACCTCCAAAGAGCTG 918
 Db 921 CCAGAAAGCAGAGATGAACTACTGGAAGCTCACTAAATTTGAATCTTTTAACTGGA 980
 QY 919 CTGACCTTCAGCCAAAGAAACAAG-----TCTTCACTGCAAGTGTGTGTGTGTGTGT 971
 Db 981 CTTATCTTGAATTCAAAAGAGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1040
 QY 972 CTTTCCTATAGACCTCTAGATCTATGTCTCAGAAAGCATTTAAAGCAATTCCTTCAT 1031
 Db 1041 TCTGCCAAAGGACCAAGAAAGATCTGTGTGAAAGATTTCAAGCACTGTCTCTCTCAT 1100
 QY 1032 CATCGGAGTCAATTAACAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11082
 Db 1101 AGTGGGCACTCAACAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1160
 QY 1083 TCTGAGATCTCAGTGTGCTCAACAAGTCCCTTGCCTCCATCTGATACAAAAC--AT 1139
 Db 1161 CGCTGAAGGCAAACTGACCAAGAAAGCAAGCAATTTCTTGTGGAAGTCTTACCCAAAC 1220
 QY 1140 CTGCAACATCCCGCTCAGTATTTGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1199
 Db 1221 ACTTAAATCTCTGAAATATGATTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280
 QY 1200 CTCCCTGACTGAATTCGAGACAGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1259
 Db 1281 TGAACCTACCAAAAAGAAAGCTGTTCAGAGACTTATGATGTGTGTGTGTGTGTGTGT 1340
 QY 1260 CCCTGCACTGATCAAGCTGATCAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1319
 Db 1341 CCATATGATGATTTGT 1400
 QY 1320 GTTTTGGCAACCGGCTCAGT 1379
 Db 1401 GTTGTGATATCTGCCCAAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1460

QY 1380 CGGTATGAATGCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1439
 Db 1461 TGGTATGAGATCTTCTCAGTATTTGATTTGATCTCATTTTAA----- 1502
 QY 1440 CGAAGAGCCAGGAGAGAGAGAGAGTACTGAGCCGGAAGATGATGAATTAATCTGGCTAC 1499
 Db 1503 AGATGTGTCTCAGAAAG 1562
 QY 1500 CTTTGCTCAACCGGAGAAATCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1559
 Db 1563 CTTTGT 1622
 QY 1560 GACTAGACATCTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1619
 Db 1623 GAAGGAAGGATATCTGAAGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1682
 QY 1620 GCGGTGTGATTTTGTGAC 1637
 Db 1683 AGAAGTGAATTTTGTGGC 1700

RESULT 10
 HSM803013
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2888)
 Pouska,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
 Mewes,H.W., Well,B., Amid,C., Oeangner,A., Fobo,G., Han,M. and
 Wiemann,S.

REFERENCE
 AUTHORS
 CONSRM
 TITLE
 JOURNAL
 COMMENT
 The German CDNA Consortium
 Direct Submission
 Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 Sequenced by DKFZ (German Cancer Research Center.
 Heidelberg/Germany) within the CDNA sequencing consortium of the
 German Genome Project. This clone (DKFZp43N0935) is available at
 the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
 Berlin, Germany. Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/c1.cgi?cloneID=DKFZp43N0935
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.

FEATURES
 source
 location/Qualifiers
 1..2888

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp43N0935"
 /db_xref="taxon:9606"
 /clone="DKFZp43N0935"
 /tissue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSport1, host
 DH10B; gates NciI + SalI"
 /dev_stage="adult"
 /note="carboxylesterase 2 isoform 2"
 1..2888
 /gene="DKFZp43N0935"
 954..2777
 /gene="DKFZp43N0935"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAD28531.1"
 /db_xref="GI:19584507"
 /db_xref="UniProt/Swiss-Prot:O00748"
 /translation="MTAQSRSPTPTFPGPSQRFLPCPVQVPRLGRALIHCHWTPG

OPLGQQRVRRORETESEPTWMLHRLARLSAVACGLLLLVRCGQDSASIRRTTH
 GDLVSLVHVKGANVOTFLGIIPKAPKPLRLRPAPREPSPMSVWRGTHPAMLC
 ODVTVSESPISOPMMTPSPDSMDSDCIYLSTYTPAHSPREPSMTPMWHGALVPG
 MASLYDSKMLALENVVVVITOTRLGVLGPEPSTGDKHATNGWITDOVALHVRQONT
 AHFGNPKDVTI FGSAGTSVSLVSTISQGLPHGALMEGVALLPBLIASADVI
 STVVNLASCDVDDEBALVGLRGSKEIILANKPFXIPEVVGVFLPRHQELIA
 SADFPVPSIVGVNNNEFGMLIPKVMRIYDQKEMDRSAQALKMLTLMLPTFG
 DLAREYIGDNDDPOTLOAQPOEMMAISHFVIALOVAHFQCSRAVYVEFOHOPSM
 LKNIRPPHMKADHVKPTREEROLSRMMVYMANPABNPNBEGLPHEMLPQOEBYL
 QNLQNPANGRALKARLOPFMKALPOKIOLEPEBRHTEL"

Query Match 19.8%; Score 345.8; DB 3; Length 2888;
 Best Local Similarity 54.8%; Pred. No. 3.6e-87;
 Matches 850; Conservative 0; Mismatches 617; Indels 84; Gaps 5;

116 GGAAACACAGGCTGGAGTGAATTCAGGGGACAGCACTGCTGCTGGGAAGCCCTGTGC 175
 1246 GGACACACACACAGGGGGCAGGTGTGGGAGTCTTGTCATGTGAAGGGCGCCAAATGCCG 1305
 176 CTGTGACGCTTCTCCGAGTCCCTTTGCTGCTCCCGGCTGGGATCCCTGGCATTTA 235
 1306 GGGTCMAACCTTCTCGGAATTCATTGGCAAGCACCTCTAGGTCCGCGCATTTTG 1365
 236 CGAACCCGCGAGCTGCATGCCCTGGGATTAATTGGGAGAGCACCTCTACCTAATTT 295
 1366 CACCCCTGAGCCCTCTGATTTTGAAGTGTGTGAGGATGGAACACCTCATCCGGCCA 1425
 296 TGTGCTCCAGAACTCAGAGTGGCTGTCTAGA--TCAGACATGCTCAAGTGCATT 352
 1426 TGTGTACAGAGACCTCAACGAGTGAAGTCAAGTTTCTTAGCCAGTTCAACATGACCT 1485
 353 ACAGCAATTCGGAAGTGTCAAGAACTGCTCTCACTGAACATCTATGCGCTCCGCCAG 412
 1486 TCCCTTCGCACTCCATGTCTGAGAACTGCTGTGATCTCAAGCACTCAACGCGGCCA 1545
 413 CGGATPACAGGCTCAAGTCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 472
 1546 GCGATGAAAGCTCTTAACTGCGGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1605
 473 GCTGACCTCCATCTTGTGATGGGTCCGCTGCTGCTGTGATGAGACGCTGTGTGTGTG 532
 1606 GGTGAGCTTCTTGTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1665
 533 TGTGTCAAGTACCGGCTGAGAAATTTGTTTCTTCAACCAATGGGATCGCATGCTCCGG 592
 1666 TCATCAGATACCGCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1725
 593 GGAACCTGGGCTTCAAGACAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652
 1726 GGAACCTGGGCTTCAAGACAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1785
 653 TCTTGGTGGGAGCCCACTGCTGTGACATCTTGTGGAGTCCGCGGAGGACATTAATGTG 712
 1786 ACTTTGAGGCAACCTGTGACGCTGTGACATTTTGTGGAGTGTGGGGTGTGGACAGTGT 1845
 713 TTTTACTTATCTGTCTGT 772
 1846 TGTCTTGT 1905
 773 GTGGGGTGGCCATCATCTTACCTGAGAGCCCATGATATAGAGAGTGTGAGACCTGTG 832
 1906 GTGGGGTGGCCCTCTGCGCGGCTCTATTTGCAAGCTCAAGCTGA-----TGTCAATTC 1959
 833 AGGTGGTGGCAATTTCTGT 892
 1960 CGGTGGTGGCAACCTGT 2019
 893 TGAAGCAAAACCTTCAGAGAGT 952
 2020 TCCGGGGCAAGAGTGAAGAGATTTCTTGAATTAACAAGCTTTCAAGATGATGATCCCG 2079
 953 GAGTGGT 1012

2080 GAGTGGT 2139
 1013 TTAAGCAATTTCTTCCATCATTCGAGTCAATTAACAGAGTGTGTGTGTGTGTGTGTGTGT 1070
 2140 TTTAGCTGT 2199
 1071 -----TATGAAGAGGCTCTGAGATCTTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1120
 2200 AGTCAATGAGATCTATGATTAACAGAGAAATGAGACAGAGAGGCTTCCAGGCTGTCTC 2259
 1121 TCCATCTGATACAAACATCTGTGACATCCGCTCTCATGATTTGTGACCTTTGTGCTAATG 1180
 2260 TCCAGAAATGTATTAACGCTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2319
 1181 AATATCTTCATGACAAAGACCTCCCTGATCTGAAATCCGAGACAGTCTTCTGTGACTTGTCT 1240
 2320 AGTACATTTGGGACATGAGGAGATCCCGACAGCCCTCCAGGAGATTCAGAGAGATGATG 2379
 1241 GAGATGT 1300
 2380 CGGATCTCATGTTTGT 2436
 1301 CACTGTCTATCTTCTATGAGTTCGGCACCGGCTCTCATGTCTTTGAAACACAGAACCGG 1360
 2437 CCGTGTGTATCTTCAAGAGTTCAGACATCAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGT 2496
 1361 CTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTTGTGTGTGTGTGTGTGTGTGTGTGT 1420
 2497 CCGACATGAGGACAGACCATGTATA----- 2521
 1421 AGGGGACATTTGTATGTTTGAAGGACCGAGGAGGAGAGTACTGAGCCGAGAG 1480
 2522 -----ATTCACTGAGGAGAGAGGACGACTTAAGCAGAAAG 2556
 1481 TGAATGAATATCTGGGCTACCTTTGTGTGAAACCGGAAATCTATAGGAAACGACTGTCTC 1540
 2557 TGAATGAATATCTGGGCTACCTTTGTGTGAAACCGGAAATCTATAGGAAACGACTGTCTC 2616
 1541 TGTGCGCAGCTTATTAATCTGACTGAGCAGTACTCTCAGCTGTGAATCATGAGCTGTG 1600
 2617 ACTGCGCGCTGTGTGACACGAGGAGCAATATCTGAGCTGTGAACCTTACGCTGTGCTGTG 2676
 1601 GACAGAGACTCAAGAAACCGGCGGTGATTTTGTGACGACGACCATTCCTCC 1651
 2677 GCGGCGCTGTGAAGGCCACAGAGCTTCAAGTTCTGTGAAGAGCGCTGCCCC 2727

RESULT 11
 HSM806270
 LOCUS HSM806270 3909 bp mRNA linear HTC 22-SEP-2004
 DEFINITION Homo sapiens mRNA; cDNA DKFZp686H0466 (from clone DKFZp686H0466).
 ACCESSION BX538086
 VERSION BX538086.1 GI:31874222
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3909)
 REFERENCES
 Bloeker H., Boeher M., Brandt P., Mewes H.W., Well B., Amid C.,
 Oanger A., Fodor G., Han M. and Wiemann S.
 The German cDNA Consortium
 Direct Submission
 Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp686H0466) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.

Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/c1.cgi?cloneid=DKFZp686H0466>
 Further information about the clone and the sequencing project is
 available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

Source

1. 3909
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="RZPD:DKFZp686H0466"
 /db_xref="taxon:9606"
 /clone="DKFZp686H0466"
 /tissue_type="endometrium carcinoma cell line"
 /clone_id="686 (synonym: hicc3). Vector pSPORT1_Sfi; host
 DH10B; sites SfiI + SfiII"
 /dev_stage="adult"
 /note="carboxylesterase 2 isoform 2"
 1. 3909
 /gene="DKFZp686H0466"
 1002. 2825
 /gene="DKFZp686H0466"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAD98009.1"
 /db_xref="GI:31874223"
 /db_xref="UniProt/Swiss-Prot:O00748"
 /translation="MTAQSRSPTTFPPSPSQRTPLPVOTPRIGKALHICWTDPG
 GLGBOQRYRORSTETSEPTMBLRLRLSLAVAGLILVLRGGQDSASIRTHHT
 GOVLGSLVHAKNAGNVORFLGIIPAKPIGLRPARPPSPSNGVDTTHPMCL
 DLTAVESFSLQPNMTFSPSDMSDECLYITTPHSHSGNLVLPVAILHGAALVFG
 MASLIDGSLALENVVVIQIRGLVLFSTGDAHAGNNGLDPAVALHAWQONI
 AHFGPDRVTI FGSAGTSVSLVSPISQGLPHGALMEGVALLPGLIASADVI
 STVFNLSACDVIDSALGVCLRGSKSEIILAIINKPKMIPGVVGVFLPRPOLLA
 SADPOVPSIVGVNNNEFGMLIPKVMRIYDTQKEMDRASQAALKMLTLMLPTSG
 DLTREYIDNGDPOTLQAFQFQEMADSMFVLPALQVHFOCSRAPVYFEEQHPSM
 LKNIRPHMKADHVKTREERDLSRKMKCYMANPRANKPNSEGLPHMPLFQESQYL
 RLNLDPAYGRALKAHRLQFMKALPQKIQELEPERHTEL"

gene

CDS

116 GGAACACGAGCTGGATGATTCAGGGCAGCAAGTCACTGTGTGGAGAGCCCTGTGC 175
 1294 GGAACACACACAGCGGGGAGGTGTGGGGAGCTTGTCCATGTGAAGGCGCCAAATGCG 1353
 176 CTGTGAAGTGTCTCGGAGTCCCTTGTGTGCTCCCGCTGGAGATCCCTGCATTTA 235
 1354 GGGTCCAAACCTTCTGTGGAAATTCATTGTGCCAAGCCACTCTTAGTCCGCTGCATTG 1413
 236 CGAACCCGACGCTGCATCGCCCTGGAGTAACTTGCGAAGCAGCTCTACCTTAATT 295
 1414 CACCCCTGAGAGCCCTGAAATCTTGAGTGGGTGAGGATGGAACCAACCATCCGGCA 1473
 296 TGTGCTTCAAGAACTCAAGTGGCTCTTTAGA---TCAACACATGCTCAAGGTGCAAT 352
 1474 TGTGTCTACAGAGACTCAACCGCAGTGAAGTCAAGTTTCTTAGCAGTTCAACATGACCT 1533
 353 ACCGAAATTCGAGAGTTCAGAAAGCTGCTCTACCTGAACATCATATGCGGCTCCAGC 412
 1534 TCCCTTCGAGTCCATGTCTGAGAGTCCCTGTACTTACGATCTCAACGCGCGCCATA 1593
 413 CCGATACAGGCTCAAGCTCCCGCTTGTGTGTGGTTCCTCCAGAGAGTCCCTTCAAGACTG 472
 1594 GCCATGAAGGCTCAACCTGCGGATGATGTGTGATCAAGGTGTGCGCTGTGTTTTTG 1653
 473 GCTAAGCTCATCTTTGATGAGGTCCCGCTGGCTGCTTATGAGAGAGTGTGTTGTGG 532
 1654 GCATGGCTTCTTGTATGATGATGTTCAATGCTGCTGTGGAAAGAGTGTGTGTGTCA 1713
 533 TCGTCAAGTACCGGCTAGAAATATTTGGTTTCTTCAACAACATGGAGATCAGATGCTCGG 592
 1714 TCATTCAGTACCGGCTGGGTCTCTGGGCTTTCTTCAAGCATGAGAGCAAGCAACGAAACG 1773

ORIGIN

Query Match 19.7%; Score 344.2; DB 3; Length 3909;
 Best Local Similarity 54.7%; Pred. No. 1.1e-86;
 Matches 849; Conservative 0; Mismatches 618; Indels 84; Gaps 5;

116 GGAACACGAGCTGGATGATTCAGGGCAGCAAGTCACTGTGTGGAGAGCCCTGTGC 175
 1294 GGAACACACACAGCGGGGAGGTGTGGGGAGCTTGTCCATGTGAAGGCGCCAAATGCG 1353
 176 CTGTGAAGTGTCTCGGAGTCCCTTGTGTGCTCCCGCTGGAGATCCCTGCATTTA 235
 1354 GGGTCCAAACCTTCTGTGGAAATTCATTGTGCCAAGCCACTCTTAGTCCGCTGCATTG 1413
 236 CGAACCCGACGCTGCATCGCCCTGGAGTAACTTGCGAAGCAGCTCTACCTTAATT 295
 1414 CACCCCTGAGAGCCCTGAAATCTTGAGTGGGTGAGGATGGAACCAACCATCCGGCA 1473
 296 TGTGCTTCAAGAACTCAAGTGGCTCTTTAGA---TCAACACATGCTCAAGGTGCAAT 352
 1474 TGTGTCTACAGAGACTCAACCGCAGTGAAGTCAAGTTTCTTAGCAGTTCAACATGACCT 1533
 353 ACCGAAATTCGAGAGTTCAGAAAGCTGCTCTACCTGAACATCATATGCGGCTCCAGC 412
 1534 TCCCTTCGAGTCCATGTCTGAGAGTCCCTGTACTTACGATCTCAACGCGCGCCATA 1593
 413 CCGATACAGGCTCAAGCTCCCGCTTGTGTGTGGTTCCTCCAGAGAGTCCCTTCAAGACTG 472
 1594 GCCATGAAGGCTCAACCTGCGGATGATGTGTGATCAAGGTGTGCGCTGTGTTTTTG 1653
 473 GCTAAGCTCATCTTTGATGAGGTCCCGCTGGCTGCTTATGAGAGAGTGTGTTGTGG 532
 1654 GCATGGCTTCTTGTATGATGATGTTCAATGCTGCTGTGGAAAGAGTGTGTGTGTCA 1713
 533 TCGTCAAGTACCGGCTAGAAATATTTGGTTTCTTCAACAACATGGAGATCAGATGCTCGG 592
 1714 TCATTCAGTACCGGCTGGGTCTCTGGGCTTTCTTCAAGCATGAGAGCAAGCAACGAAACG 1773

593 GGAACCTGGGCTTCAAGAGCAGAGTGGCTGTCTGTCTGTGGTCCAGAAATCATGAGT 652
 1774 GCAACTGGGGCTACCTGAGCAAGTGGCTGCACTACGCTGGGTCCAGCAAGATATCGCC 1833
 653 TCTTCGTTGGGAGACCCCAAGCTCTGAGACATCTTTGGGAGTCCGGGAGCCATTAAGT 712
 1834 ACTTGGAGGCAACCTTGACCGGTGACCATTTTGGGAGTCTCGGGTGGCAGAGTGG 1893
 713 TTTTACTTTTACTGTCTCCATAGCCCAAGGCTTATTCAGAAAGCATCATGAGA 772
 1894 TGTCTTCGTTTGTGTGTCCTCCCATATCCAGAGACTCTTCCAGGAGCATCATGAGA 1953
 773 GTGGGGTGCATCATCTCTTAACCTGAGGCCCATGATTAAGAAAGTGAAGTGCCTGC 832
 1954 GTGGGCTGGCTCTCTGCGCGGCTCATTTGCAAGCTCAGCTGA-----TGTCACTCCA 2007
 833 AGGTGGTGCATTTCTGTGTACATATGCTCAGATCTTGAAGCCCTGCTGAGTGC 892
 2008 CGGTGTGGCCAACTGTCTGTGTGACCAAGTTGACTTGAAGGCCCTGTGGTGGCTGCC 2067
 893 TGAGACAAAACCTTCAAGAGTGTGACCTCAGCCAGAAAACAAAGTCTTTCACTC 952
 2068 TCGGGGCAAGATGAAGAGATTTCTGCAATTAACAAGCTTTCAAGATGATCCCCG 2127
 953 GAGTGTGATGTGCTTCTTCTTAATGAGCTTAAGATCTATTGTCTGAGAAAGCAT 1012
 2128 GAGTGTGATGTGGGTCTTCTCTGCGCCAGGACCCCAAGAGACTGCTGCTGCCACT 2187
 1013 TTAAGCAATCTCTTCCATCATCGAGATCATTAACAAGATGTGGTCTCTGCTGCC-- 1070
 2188 TTCAGCTGTCTCTAGCATTTGTGTGTCACCAACAAATGATGGCTGCTACCCCA 2247
 1071 -----TATGAAGAGGCTCTGAGATCTCAAGTGTCCCAAGTCCCTTGC 1120
 2248 AGGTCAATGAGATCTATGATACCAAGAAATGACAAGAGAGCCCTCCAGGCTGCTC 2307
 1121 TCCATCTGATCAAAAACATCTGCAATCCGCTCAGTATTTGCACTTGTGGCTAATG 1180
 2308 TGCAGAAATGTTAAGCTGTGATGTGTCTCTTACATCTGTGAGACTGCTGAGGAGG 2367
 1181 AATCTTCAATGACAAACATCTCTGACATCCGCTCAGTATTTGCACTTGTGGCTAATG 1240
 2368 AGTACATTTGGGAGCAATGGGGATCCCAAGACCTTCAAGGCACTTCAAGAGATGATGG 2427
 1241 GAGATGTGTTCTTGTGTGCTCTGCACTGATCAAGCTGATATCAAGAGATGCTGATG 1300
 2428 CGGACTCATGTTGTGATCTCTGCACTCAAGATGACATTTCAAGT---TTCGCGGG 2484
 1301 CACCTGTCTCTTGAAGATTTTGGGACCGGCTCAGTGTCTTGAAGACAGAGCCGG 1360
 2485 CCGCTGTGATCTTCAAGATTTCAAGATCAAGCTGAGCTGTCAAGAAATCATGAGGCAC 2544
 1361 CTTTGTCAAGGCGACCAAGCTGATGAAGTCCGCTTGTGTGCGGTGTGCTTCTGTA 1420
 2545 CGCAATGAAGGAGCAACATGTTA-----ATTCATGAGGAAGGAGCAGCTAAGCAGGAAGA 2569
 1421 AGGGGACATTTGTATGTTGAAAGAGCCACGAGAGAGAAATTAATGAGCCGGAAGA 1480
 2570 -----ATTCATGAGGAAGGAGCAGCTAAGCAGGAAGA 2604
 1481 TGATGAATATCTGGGCTACCTTGTGCAACCGGGAATCTTAATGGAAAGCACTGTCTC 1540
 2605 TGATGAAGTATCGGCGCACTTGTGCAAGAAATGGAAACCAATAGCGAGGTCTCCAC 2664
 1541 TGTGGCAAGTTAATCTGACTGACAGTACCTTCAAGCTGACTTGAACATGAGGCTGG 1600
 2665 ACTGGCGCTGTTCGACCAAGAGAGCAATACCTGTGGCTGAAACCTACAGCTGCGGTGG 2724
 1601 GACAGAGATCTCAAGAAACCGCGGTGATTTTGAACAGACATATCCCC 1651
 2725 GCGGGCTCTGAAGGCCCAAGGCTTCAAGTCTGAAAGAGGCGGCTGCC 2775

RESULT 12
 CP111083 1918 bp mRNA linear EST 23-JUL-2003
 LOCUS Shultzomicron4334 Rat lung airway and parenchyma cDNA libraries
 DEFINITION Rattus norvegicus cDNA clone Contig3845 5', mRNA sequence.
 ACCESSION CP111083
 VERSION CP111083.1 GI:33167595
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus
 1 (bases 1 to 1918)
 Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
 Padua,A.M., Gureke,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
 Plopper,C.G. and Buckpitt,A.R.
 Gene expression analysis in response to lung toxicants: I.
 Sequencing and microarray development
 Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
 JOURNAL
 COMMENT Contact: Shultz MA
 Dept. of Molecular Biosciences, School of Veterinary Medicine
 University of California, Davis
 1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
 Tel: 530 752 0793
 Fax: 530 752 4698
 Email: mshultz@ucdavis.edu
 Average Phred score is 20 or better. All poor quality data (Phred <
 20) and vector/linker sequence has been removed.
 High quality sequence stop: 1918.
 FEATURES
 Source
 1..1918
 location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="Contig3845"
 /sex="male"
 /tissue_type="airway or parenchyma"
 /dev_stage="adult"
 /clone_lib="Rat lung airway and parenchyma cDNA libraries"
 /note="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
 Site 2: Not I; mRNA was isolated from microdissected rat
 lung airways and parenchyma tissues."
 ORIGIN
 Query Match 19.4%; Score 338; DB 7; Length 1918;
 Best Local Similarity 54.2%; Pred. No. 5.5e-85;
 Matches 869; Conservative 0; Mismatches 660; Indels 75; Gaps 6;
 Oy 91 GGGCTTGTGCTGAGGGCCACAGAGAACACAGGCTGGAGTGTTCAGGGCAAGCAA 150
 Db 83 GGGATTCATCTCTCAACCTGTGGGAACTGTTAAAGCAAGTCTGGGGAAGTAT 142
 Oy 151 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTCTTCTCGAGTCCCTTTGTGCT 210
 Db 143 GTCAATTGGAGGATTTGCACAGGCTGTGCTGTTTCTCGGAATCCCTTCGCAAG 202
 Oy 211 CCCCCGTGGATTCCTCGATTTTACGAACCGGACGCTGCATGCGCTGGATTAAGTTC 270
 Db 203 TCCCTCTTGGCTCTTTAGGTTTGTCTCAACACAGCTCGACAGTCTTGGAACTTTGTG 262
 Oy 271 CGAGAGCACTCTTACCTTAATTTGTGCTTCAGAACTC-----AGAGTGGCTGCTC 324
 Db 263 AAGAAATTAATCTCTTACCTTACCTTAATTTGTGCTTCAGAACTGTTGGAGGCGAGTTCTC 322
 Oy 325 TTGATCAACAGATGCTCAAGTCAATTACCGAAATTTGGAGTGTCAAGAAATCTGCTTC 384
 Db 323 TCAGAGCTTTTACCAACAGGAAGAAACATTTCTTTACAGTTTTCGAAACATGCTTC 382
 Oy 385 TACCTGAACATCTATGCGCTGCCCAAGCGGATACAGGCTCCCGCTCTTGTG 444
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 383 TACCTGAATGTTTATATCTCCGCTGACTTGACAAAGAAACAGCGGTTACAGTATGCTG 442
 Oy 445 TGGTCCAGAGAGGAGGCTTCAAGCTGAGCTCAGGCTCCATCTTATGATGGTCCGCTG 504
 Db 443 TGGATTCATGAGAGGAGGAGCTGATGATGATGATATCACTTATGATGACAGGCTTC 502
 Oy 505 GCTGCTATGAGAGAGCTGCTGTTGTGCTGCTCAATACCGGCTAGAAATTTGGTTTC 564
 Db 503 TCTGCCATGAAAAATGTGTGTGTGTGACATTCAGTATCGCTTGGCATCTGGGATTC 562
 Oy 565 TTACCAACATGGGATCAGATGCTCCGGGAACTGGGCTTCAAGACCAAGTGTGCT 624
 Db 563 TTCAGACAGGGGATGAACACAGCCAGGGCAACTGGGGTCAATTTGACCAAGCTGTGCA 622
 Oy 625 CTGTCCTGGTCCAGAAACATGAGTCTTGGTGGGAGCCCAAGCTGTGCAATC 684
 Db 623 CTGACCTGGTCCAGACAAATTTCCAACTTGGGGCAACCAAGGCTGTGACATC 682
 Oy 685 TTTGGCAGTCCGCGGAGCCATTAAGTCTTCTACTTATCTGTCTCCATGGCCAA 744
 Db 683 TTTGAGATCTGACAGAGGTTTCACTGTCTCTGTTCTTGTATCTCTCGCCAA 742
 Oy 745 GCTTATTTCCAAAGCCATCATGAGAGTGGGCTGGCCATATCTCTTACTGAGGCC 804
 Db 743 AACCTTTCCACAGGGCCATTTCTGAGAGTGTGTGCTCCATATATCTGTATTA 802
 Oy 805 CATGATTTATGAAAGAGAGGAGCCTGACAGGCTGTGCAATTTCTGTGTAAATGCG 864
 Db 803 ACAGATTCAGAGCCCATGCTATATCTGATGTTACTCTTCTGGGTAAACCAACACA 862
 Oy 865 TCAGACTCTGAGGCCCTGCTGAGTGTGCTGAGGACA----- 900
 Db 863 TCAGCTGTATGTTTCAATTTGCTGCGCCAGAAAGAGATGAACTCTGAGACTTCA 922
 Oy 901 -----AACCTTCAGAGAGCTGTGACCTTCAGCCAGAAACAAAGTCT----- 945
 Db 923 TTAATATGAATCTTTTAACTGACCTATTTGGAAACCCAAAGAGCTATCTCTTC 982
 Oy 946 TTCACTCAGAGTGTGATGAGTCTTCTTCTATATGAGCCCTGTGATCTATTTGTCTAG 1005
 Db 983 CTACTACTGTGATGATGAGTGTGTGCTGTCAAAAGACACAGAAAGATCTGTGCTGAG 1042
 Oy 1006 AAGGATTTAAAGCAATCTCTTCATCATGCGAGTCAATTAACACAGAGTGGCTTCTG 1065
 Db 1043 AAGATTTCAACAGGCTCCCTCACTATGAGGATCAACAGCAAGATTTGGCTGATC 1102
 Oy 1066 CTGCTATG-----AAGAGGCTCTGAGATCTCACTGAGCTTCAACAGTCCCT 1115
 Db 1103 ATTCAAATGCTTATGAGGCTATCACTTCCGAAGCAAACTGACCAAGAAACAGCCAAT 1162
 Oy 1116 TGGCTTCATCTGATACAAAC--ATCTGCAATCCGCGCTCAGTATTTGCACTTGTG 1173
 Db 1163 TCCCTCTGTGGAATCTTACCAACACTGAAATCTCTGAGAAATGATTCATAGGCT 1222
 Oy 1174 GCTATGATATCTTCATGACAGACATCCCTGATCGAATTCGAGACAGTCTTCTGAGC 1233
 Db 1223 TCTGAAATATTTAGAGGAGCAATGACCTGCGCAAGAAAGACCTGTTCAAGAGC 1282
 Oy 1283 TTGTTGAGATGATATTTGTGTGCTCCATCAGATGATGTGTCTGAAGTCAAGAGAT 1342
 Db 1294 GGTGTGACCTGTCTACTTCTATAGTTCGAGACCGGCTCAGTGTCTTGAAGACG 1353
 Oy 1343 GCTGAGGCCCACTTATGATATGATTTGATGCGCCAAAGCTTGTATACAGCAAG 1402
 Db 1354 AAGCGGCTTTTGTCAAGCCAGACAGCTGATGAAGTCCGCTTGTGTTCGATGATGCGC 1413
 Oy 1403 AGGCCAAGACAGATATGAGAACATGTATGAATCTTCTCAATTTGGATCTCA 1462
 Db 1414 TTCTGAAGGGGAGACATTTATATTTGAAAGAGCCACGAGAGAGAAATTTACTGAGC 1473
 Oy 1463 TTTTAA-----AGATGGCGCTCAGAAAGAGAGCAATCTCAGC 1504
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY	1474	CGGAAGATGATGAAATATCTGGGCTTACCTTTGCTCGAACCGGGAATCTTAATGGGAACGAC	1533
Db	1505	AAATATGTATGTAATAATCTGGGGCAACTTTGCTCGAATGGGAACCTTAATGGGGAGGG	1564
QY	1534	CTGTCTCTGGGGCAGACTTATATCTGACTGAGCAGTACCTCAGGTGGACTTTGAACATG	1593
Db	1565	CTGGCCCATTTGGCCAAATATATGACCAAGGAAGAGTTACCTTAAGATTTGTGCTCAACT	1624
QY	1594	AGCCTCGACACAGACTCAAAGAACCGCGGTGGATTTTGGAC	1637
Db	1625	CAGGCAGCCAGAGGCTGAAGGACAAAGAAAGTGGCTTTTGGTTC	1668
RESULT 13			
LOCUS	BC019926	2038 bp	linear HTC 20-SEP-2002
DEFINITION	Mus musculus, clone IMAGE:5123923, mRNA.		
ACCESSION	BC019926		
VERSION	BC019926.1 GI:18044766		
KEYWORDS	HTC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2038) Strausberg, R. Direct Submission Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REFERENCE	NIH-MGC Project URL: http://mgc.nci.nih.gov		
AUTHORS	Contact: MGC help desk		
TITLE	Email: cgapbs-remail.nih.gov		
JOURNAL	Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunnarsson, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.		
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov Series: IRAX Plate: 42 Row: P Column: 20 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein This clone has the following problem: frame shifted.		
COMMENT	Location/Qualifiers 1..2038		
FEATURES	Source		
SOURCE	/organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:5123923" /tissue_type="Liver, normal, 5 month old male mouse." /clone_id="NCI CGAP_L19" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"		
ORIGIN			
Query Match	19.0%; Score 331.4; DB 3; Length 2038;		
Best Local Similarity	53.1%; Pred. No. 4,4e-83;		
Matches 839; Conservative	0; Mismatches 711; Indels 31; Gaps 5;		
QY	88	ACTGGGCTTCTGCTAAGGCCACAGAGAACACCAAGGCTGGAGATGATTCAGGGCAAG	147

D	98	ACTGAGACCCAAAGTCATTACGCTCGAAGTGGATACCCCTGGGTCTGTTCCAGGCCGG	15
Q	148	CAAGTCATGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCTCGAGTCCCTTTGCT	207
D	158	CAGGTGGTGTGAAGGACACAGACCGCATGTAATGTCTTCTCGGGCATCCCTTTGCT	217
Q	208	GCTCCCCCGGTGGAGTATCCCTGGATTATAGAACCCCGAGCTGATCGCCCTGGATTAC	267
D	218	CAAGCACAGATGGAGACTTTCCTGGTTCTACGTCACACTCCACACAGCCCTGGAAAGT	277
Q	268	TTGCGAAGAACCACTTCCTAACCCCTAAATTTGTGCTCCAGAACTCAGAGTGGCTCTTA	327
D	278	GTGAGAGATGCCAGCATCAATCCCAATGTGCTTCAGAGATGTAGAAATATCATAC	337
Q	328	GATCAACATAGTCTCAAGTGTACTTACCCGAAATTTGGAATGTCAAGATCTGCTTAC	387
D	338	AGCAATTTTGGCTTCATGAAATTAATTTCCATTTCTAGAGACTGCTCGAAC	397
Q	388	CTGAACATCTATGCGCTGCTCCACGCGGATACAGGCTCCAAAGCTCCCGCTTGGTGG	447
D	398	CTCAACATCTACAGCCCACTGATGATCATGCAAGGGAACTAAAGCCGCTCATGTATGG	457
Q	448	TTCCACAGAGTGCCTTCAAGACTGGCTCAGCTCATCTTGTATGATGGTCCGCTGCT	507
D	458	ATCCACGAGGCTCTCTGCTGTGGTCTCTCCACATCTCAGGATGATCAGACTGGCC	517
Q	508	GCTATGAGAGCGTGTGTTGTGTGTCGACGATACCGGCTTGAATATTTGTTCTTC	567
D	518	GCGTAGGGGAAATGTGTATGTGTCTACTGTCCAGTATCGCTTGGGATCTTGGCTTCTC	577
Q	568	ACCAATGGGAATCAGATGCTCCGGGGAATCGGCTTCAGAGACAGGCTGGCTGTG	627
D	578	AGCATCTGAGACAAACACATGCAAGCAACAGGGATTCCTGGATGTGGTGTGCTCTT	637
Q	628	TGCTGGTCCAAABAACAATCGAGTTCTTCGGTGGG-GACCCAGCTCTGTACCATCTT	686
D	638	CGCTGGGTTCAGGGAAACATACCCCCCTTTGGGGGTGATTCCAACTGTGTCAATATCTT	697
Q	687	TGCGAGTCCGCGGAGCCATAGTGTCTTATGCTTATATCTGTCCCATGCGCAAGG	746
D	698	CGGTAACTGTGTGAGGATATTTGTCTACCTCCTGTTCTGTCTTCAATATCTGTGG	757
Q	747	CTTATTCACAAGGCATCATGAGAGTGGGGTGGCCATCATCCCTTACTGAGAGCCCA	806
D	758	GCTCTTCCACAGAGCATATCCCAAGTGGGATGTCAAGCAATTAATGATGG-----A	811
Q	807	TGATTATGAGAAAGATGAGACCTTGACAGTGGTTGCAATTTCTGTGTAAACAATGCTTC	866
D	812	AGACATGAACCATGTGCTGAAGCTCAGAACTTTGCCAATTTCTGTGGCTGTGCTGCTG	871
Q	867	AGACTCTAGAGCCCTGTGAGGTGCTGAGGACAAACCTTCAGAGAGCTGTGACCT	926
D	872	ATCCCACTAGCTGAGTGTCCAGTGTGTTCTCAGAAAGAGAAAGACCTT-----AT	925
Q	927	CAGCAGAAACAAGTCTTCACTCGAGTGTGTGATGTGTGTTCTTTCTTAATGAGCC	986
D	926	CAGCAGAAACAAGTGAATTTCTTAATGTCAATATCTGTGATTAAGTTGAGAACTCT	985
Q	987	TCTAGATCTATGTCTCAGAAAGATTTAAAGCAATTCCTCAATCTCGAGTCAATTA	1046
D	986	AGAGAGTCTTAAGACACGCAATTCACAGCTGTGCCCTTCTTTGGAGTCAACAA	1045
Q	1047	CCAGAGTGTGGCTTCTGCTG-----TGCCTATGAGAGAGGCTCTGTAGATCTT	1094
D	1046	CCAATGAATTTGGCTGGCTTCTTAATCAATCTTGAATATCTGTGATTAAGTTGAGAACTCT	1105
Q	1095	CAGTGGCTCCAAAGTCCCTTGGCTTCATCTGTATCAAAACATCTCGACATCCCGCC	1154
D	1106	GAGCCGGGAAAGACTGTGTGAGATTTACAGGCCCTTCTTGGCCATTAATGAGAGTCCCC	1165
Q	1155	TCAGTATTTGACCTTGTGGCTAATGAATATTCCTCATGACAGACATCCCTGATCTGAAT	1214
D	1166	TGAGATATCCCACTGTATCATGATATCTGTGACAAAGGCTCAGATCAATGAGTAC	1225

QY 1215 CCGAGACAGTCTTCTGAGCTTGGAGATGTTCTTGTGGTCCCTGCACTGATCAC 1274
 DB 1226 AAGGTAATGCTTCAGAGATTCCTGGATATCTCTTCATCATCTTCACTGAACTT 1285
 QY 1275 AGCTGATATCAAGAGATGCTGTGACCTGTCTATCTTATGAGTTTGGCACCAGCC 1334
 DB 1286 CTCAAAATACCTTCAGATGCTGGGTGCTCTGTTTCTGTAGAGTTCCAGATACAC 1345
 QY 1335 TCAGTCTTGAAGACAGAACCCGCTTTTGTCAAGCCGACAGCTGATGAATCCG 1394
 DB 1346 CAGTCTTTTGAAGATTCAGCCAGCTGGGTGAAGGCTGACATGCTGTGAGAAATTC 1405
 QY 1395 CTTTGTGTGGTGGTGGCTTCTCTGA-----AGGGGACATTTGTTATGTTGAAGAGC 1448
 DB 1406 CTTTGTGTGGAGGTCTTCTCTCACTGATGAGATTCCTTGTGGCTTCCAGAGGC 1465
 QY 1449 CACGAGAGAGAGAAATTAAGTACGACCGAGAGATGATGAATACTGGGCTACCTTTGCTG 1508
 DB 1466 CACAGAGAGAGAGAGAGCTGAGCTGACATGATGAGCCCAATGAGGACAGTTTGACAG 1525
 QY 1509 AACGGGAAATCTTAATGAGAAAGACCTGTCTGTGGCCAGCTTAAATCTGATGAGCA 1568
 DB 1536 CACAGAGAAACCCCAATGGCAGAGGGGCTGCTCTTGGCCCAATTAACCACTTGAAGACA 1585
 QY 1569 GTACTCCAGCTGAGCTTGAATGATGAGCTCGGACAGAGACTCAAGAACCGGGGTGGA 1628
 DB 1586 AACTGTGAGATTTGCTTGAATCAACGAGCTGGGGTGAAGCTAAAGAGGGTGGCTACA 1645
 QY 1629 TTTTGTGACACAGACCATCCC 1649
 DB 1646 GTTCTGACAGAGACACTGCC 1666

RESULT 14
 LOCUS CR618303 1826 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CSDB014YC12 of Placenta of Homo sapiens
 (human).

ACCESSION CR618303.1 GI:50499110
 VERSION HTC; CDSLT cDNA.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 JOURNAL Full-length cDNA libraries and normalization
 REMARK Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1826)

REFERENCE 2
 AUTHORS Direct Submission
 TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 Location/Qualifiers

1..1826
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSDB014YC12"
 /issue_type="Placenta"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 17.1%; Score 298.4; DB 3; Length 1826;
 Best Local Similarity 57.1%; Pred. No. 1.2e-73;
 Matches 564; Conservative 0; Mismatches 422; Indels 3; Gaps 1;

QY 88 ACTGGGCTTCTGCTGAAGGGCCACAGAGAACACACAGCTGGATGATTCAGAGCGAAG 147
 DB 84 ACTGGGCGGAAGTTGCTCAGCTGAAGTAGACACACCTGGGTGCTGTGCGAGGCCGG 143
 QY 148 CAAGTCACTGTCTGGGAAGCCCTGTGCTGGAAGGTTCCTGGAAGTCCCTTGTGT 207
 DB 144 CAGGTGGGGGTGAAGGGACAGACCGCTTGTGAATGTCTTCTGGCATTCATTTGGC 203
 QY 208 GCTCCCGCTGGATATCCCTGGAATTTACAAACCCGACCTGATACGCTGGATTAAC 267
 DB 204 CAGCGCCACTGGGCGCTGACCGGTTCTAGGCCCAACACAGACAGACGCTGGAGGTT 263
 QY 268 TTGGAGAAAGCCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTCTTAA 327
 DB 264 GTGGGGATGCGCAGACATGCGCCCAATGTGCTTAAGAGCTGGAAGCATGAACAGC 323
 QY 328 GATCAACATGCTCAAGTGTATACCGGAATTCGAGAGTCAAGAAATCTGCTCTAC 387
 DB 334 AGCAGATTTGTCTCAAGGAAACAGAGATCTTCTCGTTTCAGAGGACCTGCTGTC 383
 QY 388 CTGAACATCTATGGGCTGCGCCAGCGCATACAGGCTCCAGGCTCCCGCTTGTGTGG 447
 DB 384 CTCAAGCTTATAGCCACAGCTGAGGTTCCCGAGGATCGGTAGGCGGCTATGATGG 443
 QY 448 TTCCAGAGAGTGCCTTCAAGACTGCTCAGGCTCATCTTGTGATGGTCCGCTGCT 507
 DB 444 GTCAATGAGAGGCGCTCTGATTAATGGCGTGCACCTCTAAGATGATCAGCTGTGGT 503
 QY 508 GCTTATGAGAGAGTCTGTGTTGTGCTCCAGTACCGGCTAAGAAATTTGCTTCTTC 567
 DB 504 GCTTATGAGAGAGTGTGTGTTGTGCTCCAGTACCGGCTTGTGCTTCTTC 563
 QY 568 ACCAATGAGAGATCAGCATGCTCCGGGGAACGTGGGCTTCAAGGACAGGTGCTGCTG 627
 DB 564 AGCACTGAGAGATCAGCATGCTCCGGGGAACGTGGGCTTCAAGTGTGCTGCTGCTG 623
 QY 628 TCTGGGTCCAGAGAGATCAGATGTTCTTGTGGTGGGACCCAGCTGTGACATCTTT 687
 DB 624 CGTGGGTCCAGAGAGATCAGATGTTCTTGTGGTGGGACCCAGCTGTGACATCTTT 683
 QY 688 GCGAGTCCGCGGAGCCATAGTGTCTTGTAGTCTTATATCTGTCTCCATGAGCCAAAGC 747
 DB 684 GGTGATCTGCGGGTGGAGCATCATCTGTGCTGTGCTGTCTCCAGTGTGCTGAGGG 743
 QY 748 TTAATTCACAAAGCATATGAGAGTGGGGTGGCCATCATCTTACCTGAGGCCCAT 807
 DB 744 CTGTTCACAGAGCCATCACAGAGTGGGATCATCACCCAGGATCATCATCTT 803
 QY 808 GATTATGAGAGAGTGAAGACTGAGGTGTGATGACATTTCTGTGTAAACATGCTCA 867
 DB 804 CACCTCTTGGCCCTTGTGATGAGAAATGGAACACCTTGGCTGACGCTCCAGCTCC 860
 QY 868 GACTTGAAGCCCTGCTGATAGTGTCTGAGAGACAAACCTCCAGAGAGTGTGACCTTC 927
 DB 864 CCGGCTGAGATGTGCTCACTGTCTTCAAGCAAGAAAGAGAGAGACTGTCTCTTGAAG 920
 QY 928 AGCCAGAAACAAATCTTCACTGAGAGTGTGATGAGTGTCTTCTTCTTAATAGCTT 987
 DB 921 AAGCTGAATAATCTATCTATCTTCAACCTGTATGAGCATGTCTTCCCAAAAGCCC 980
 QY 988 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATCTTCATCATGAGAGTCAATAC 1047
 DB 981 AAGGAACTCTGAGAGAGAGAGCCCTTCACTGTGTGCGCTTCTCATGGGTCAACAAC 1040
 QY 1048 CACGAGTGTGCTTCTGCTGCTATGA 1075
 DB 1041 CATGAGTTCAAGCTGCTCATCCTCCAGGA 1068

```

RESULT 15
CR641368      1859 bp      mRNA      linear      HTC 18-AUG-2004
LOCUS      CR641368
DEFINITION      Tetraodon nigroviridis full-length cDNA.
ACCESSION      CR641368
VERSION      CR641368.1 GI:51137813
KEYWORDS      HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE      Tetraodon nigroviridis
ORGANISM      Tetraodon nigroviridis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodontidae; Tetraodon.
              1 (bases 1 to 1859)
REFERENCE      Genoscope.
AUTHORS      Direct Submission
TITLE      Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL      : 2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
              (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      The sequences are based on single pass reads.
              More information available at
              http://www.genoscope.cns.fr/tetraodon.
FEATURES
  source      location/Qualifiers
              1..1859
              /organism="Tetraodon nigroviridis"
              /mol_type="mRNA"
              /db_xref="taxon:99883"
              /label_type="Liver"
ORIGIN
Query Match      16.8%; Score 293.6; DB 3; Length 1859;
Best Local Similarity 51.9%; Pred. No. 3e-72;
Matches 771; Conservative 0; Mismatches 694; Indels 21; Gaps 4;

104 AAGGCCACAGAGAAACACGCTGGAGTGGATTGATTCAGGGCAAGCATCTGTGCTGG 163
108 ATGCACCTGAAATCCATACCAACCTGGGAGCTGAGAGGTGATATGACAGCCGTGAAGG 167
164 GAAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCCCTTTGCTGCTCCCGCTGGG-- 221
168 GGAAGAAGCGTGGAGATCCATGATACCTGGGTGCTCCCTTGTCTAAGCACTCTGGGTC 227
222 -ATCCCTGCGATTTCGAACCCGCAAGCTGATCGCCCTGGAGATACTTGGAGAGCCA 280
228 CTGCTCTAGGCTGGCTCCACCCCGCTGTAGAGGCTTGGGAAGAGATGAGATGCTA 287
281 CCTCTACCTTAATTGTGCGCTCCAGAACTCAGAGTGGCTGCTTATATCAACATGC 340
288 CCAAGCAACCGCTGATGTGTGTTTCAGATTAGAGTTGCAACACAAATTATTAAGAAC 347
341 TCAAGTGCAATTAACCGAAATTCGAGTGTGAGAACTGCTCTACCTGAACATCTATG 400
348 TTAATTAGAGTCAACCTTCCAGACATTTCAAGAAAGATGCTTTACTCAACATTTACA 407
401 CGCCTGCCCAAGCCGATACAGGCTTCAAGCTCCCGCTTGTGTGTGTTCCAGAGGTG 460
408 CTCTGTCCAAAGACCTGAGAAATACCAAGCTCCCAAGTATGTGTGATCCAGGTGAG 467
461 CTTCAAGACTGCTCAGACCTCCATCTTTGATGGGTCCGCGCTGGCCCTGAGAGAGC 520
468 GGTGTGTTTAAAGTTCGGCTTCAATGTATGTGCTCTGCGCTGAGCTTACCAAGTG 527
521 TGTGTTGTGTGTGCTGCAAGTACCGGCTAGGAATATTTGTTTCTTCAACATGGGATC 580
528 TTGTTGTGTGTGTATCCAGTACCGACTGGGAATATCGGGCTTTCTAGAGCATGTGATA 587
581 AGAATGCTCCGGGGAACTGGGCTTCAAGACACAGGTGCTGCTGTCTCGGCTCCAGA 640
588 AGCATGATCCAGGAACATTTGTCTTGTGATCAGGTCCAGGCTTGAATGGGTCCAGG 647
641 AGAATGATGATTTCTTGTGTGTGGAGACCCAGACTGTGACCATCTTGGGAGTCCGGG 700

```

```

Db      648 AGCATATTCAACACTTTGAGAGAGACCAAGATTATGATCCATATTTGGCGAGTCTGTG 707
QY      701 GAGCCATTAAGTGTTCCTAGTCTTATATCTGTCTCCATGAGCCAAAGCTTATCCAAAG 760
Db      708 GTGAGTGTGATATCTCTGCTGCTGCTCACTGTGCTGAGGGCTTTTCCGCCACG 767
QY      761 CCATCATGAGAGTGGGGGCGCATCATCCCTTACTGAGGCCCATGATTATGAGAA 820
Db      768 CAATTGCTGAAGGTGACCTGTGCAGATGATATATTTG-----CTAGAGTGAACCC 821
QY      821 GTAGAGACCTGAGGTGTGTGACATTTCTGTGTATCAATGCGTACACTGTAGGCC 880
Db      822 ACCAGTGTATCAAGTTTGTCCGAATGATCTGTGTGATGACTAAGACACAGAAAGA 881
QY      881 TGTGAGGTGCTGAGGACAAACCCCTCAAGAGACTGTGACCTCAGCCAGAAACAA 940
Db      882 CTGCTGATTTGATCAAGAACTGACGTGAGAAACCATTTTACCTTAACTAAGAAAG 941
QY      941 AGTCTTTCACTGAGTGTGTATGATGTGCTTCTTCTTATGAGCTCTAGATCTATTGT 1000
Db      942 AGTTAAGTATCTTTTATACCCAGTGAATATTTCTTCAAAAATACAGCAGAGATGT 1001
QY      1001 CTC--AGAAACATTTAAAGCAATTCCTTCATCATGAGTCAATTAACAGAGTGTG 1057
Db      1002 TCTTTGACATGAACTTCTCATGTGCTTCAATGACTGAGTCAATATATGAGGTG 1061
QY      1058 GCTTCCTGCTGCTATGAGAGAGCTCTGTGATCTCTCATGAGTCCCAAGTCCCTTG 1117
Db      1062 GTTGTATCTGTGCTGATCTTGTCTCTCCAACTGACAGAGGGGTGACCAAGAAC 1121
QY      1118 CCTTCATCTGATCAAAAACATCTGACATCCGCGCTGATATTTGACCTG----- 1171
Db      1122 ACGTCTGTGTTCTGTGTGTGTCTATCTGATCCAAAGATGAGAGCTTGAAGAAC 1181
QY      1172 ---TGGCTAATGAATCTTCCATGACAAAGCATCTCTGACTGAATTCGAGACAGTCTTC 1228
Db      1182 GGAATCTTAATGAATATATACAGAACTGGGGAGATCTGTGAAATACAGAGCGGTACA 1241
QY      1229 TGAAGTGTGTGAGATGTGTCTTGTGTGCTGCTGCTGATCAAGCTGATATACA 1288
Db      1242 CTGAGGTGTGTGAGACATGATATTTGTCAATCTGTGCATTAAGGCTGTATGTCTACA 1301
QY      1289 GAGATGTGTGTGACCTGTCTAATCTATGAGTTTGGGCAACCGGCTCAGTGTGTGAG 1348
Db      1302 GAGATGAGGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1361
QY      1349 ACAGGAAGCCGGCTTTTGTCAAGCCGACCAAGCTGATGAATCCGCTTGTGTGTG 1408
Db      1362 ATAAAGATCAAGCTTTGTGTGAGAGTGAACCATGAGATGAATATTTTGTGTGTGT 1421
QY      1409 GTGCTTCTGAAAGGGGACATTTGTATGTTTGAAGAGCAAGAGAGAGAGAAATAC 1468
Db      1422 TTTGTTTCAAGTTTCTCATCATCAATTAAGACAGTGTGTGCTGAAAGAGAGACAGT 1481
QY      1469 TGAACCGGAAGATGATGAATATCTGGGCTACCTTTGCTGAAACGGGAAATCTTAAGGA 1528
Db      1482 TGAACCAATCATGATGAGCTATCTGGGGCAACTTTGTGTGCAAGAGGTCTCTTAATGAG 1541
QY      1529 ACAGACTGTCTGTGTGCGCCAGCTTATATCTGACTGAGCATGCT 1574
Db      1542 AGGCTCTGTGTCACTGGCCAAAGTATGAGAAACGAGAGGATGACT 1587

```

Search completed: June 15, 2005, 02:39:45
 Job time : 5801.46 secs

Db 301 CTCGAGAACTCAGATGGCTGCTCTTAGATCAACACATGCTCAAGTGCATTAACCCGAAA 360
Qy 361 TTGGAGAGTGAAGAACTGCTCTTACTTGAACATCTATGGGCTGCCCCGCGATAGA 420
Db 361 TTGGAGAGTGAAGAACTGCTCTTACTTGAACATCTATGGGCTGCCCCGCGATAGA 420
Qy 421 GGCCTCAAGCTCCCGCTTGGTGGTCCCAAGAGTGGCTTCAAGACTGGCTCAGCC 480
Db 421 GGCCTCAAGCTCCCGCTTGGTGGTCCCAAGAGTGGCTTCAAGACTGGCTCAGCC 480
Qy 481 TCCATCTTTATAGGATCCCGCTGGCTGCTTATGAGACCTGCTGTTGTTGCTCAG 540
Db 481 TCCATCTTTATAGGATCCCGCTGGCTGCTTATGAGACCTGCTGTTGTTGCTCAG 540
Qy 541 TACCGGTAGAAATATTTGGTTCTTACCAACATGGGATGAGATGCTCCGGGAACTGG 600
Db 541 TACCGGTAGAAATATTTGGTTCTTACCAACATGGGATGAGATGCTCCGGGAACTGG 600
Qy 601 GCTTTCAAGGACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 GCTTTCAAGGACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 GGGGACCCGAGCTCTGAGACCATCTTTGGGAGTCCCGCGGAGCCATTAAGTCTTAACT 720
Db 661 GGGGACCCGAGCTCTGAGACCATCTTTGGGAGTCCCGCGGAGCCATTAAGTCTTAACT 720
Qy 721 CTTATACGTCTTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
Db 721 CTTATACGTCTTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
Qy 781 GGCATATCCCTTACCTGAGAGGCCATGATATGAGAAAGTGAAGACCTGACGGTGGTT 840
Db 781 GGCATATCCCTTACCTGAGAGGCCATGATATGAGAAAGTGAAGACCTGACGGTGGTT 840
Qy 841 GCACATTTCTGTGTAAATGAGTGCAGTCTGAGAGCCCTGAGAGTGGCTGAGAGCA 900
Db 841 GCACATTTCTGTGTAAATGAGTGCAGTCTGAGAGCCCTGAGAGTGGCTGAGAGCA 900
Qy 901 AAACCTCCCAAGAGTGTGACCTGACCTGACCCGCAAAACAAAGTCTTCACTGAGTGGTT 960
Db 901 AAACCTCCCAAGAGTGTGACCTGACCTGACCCGCAAAACAAAGTCTTCACTGAGTGGTT 960
Qy 961 GATGAGTCTTTCTTCTCTAATGAGGCTCTAGATCTAATGCTCAANAAGCATTTAAACA 1020
Db 961 GATGAGTCTTTCTTCTCTAATGAGGCTCTAGATCTAATGCTCAANAAGCATTTAAACA 1020
Qy 1021 ATTCTTCCATCATGAGTCAATAACACAGAGTGGCTTCTGCTGCTATGAGAGAG 1080
Db 1021 ATTCTTCCATCATGAGTCAATAACACAGAGTGGCTTCTGCTGCTATGAGAGAG 1080
Qy 1081 GCTCTGAGATCCCTCAAGTGCCTCCCAAAAGTCCCTTGCCTCACTGATCAAAAACATC 1140
Db 1081 GCTCTGAGATCCCTCAAGTGCCTCCCAAAAGTCCCTTGCCTCACTGATCAAAAACATC 1140
Qy 1141 CTGACATCCCGCTCAGTATTTGCACTTGTGGCTAATGAATCTTCCATGACAAAGAC 1200
Db 1141 CTGACATCCCGCTCAGTATTTGCACTTGTGGCTAATGAATCTTCCATGACAAAGAC 1200
Qy 1201 TCCCTGACTGAATCCGAGACAGTCTTGTGAGCTTGTGAGATGTTCTTGTGTGCTC 1260
Db 1201 TCCCTGACTGAATCCGAGACAGTCTTGTGAGCTTGTGAGATGTTCTTGTGTGCTC 1260
Qy 1261 CTTGCACTGATCAAGCTCGATATCAGAGAGATGCTGTGCACTGCTTCTATAGAG 1320
Db 1261 CTTGCACTGATCAAGCTCGATATCAGAGAGATGCTGTGCACTGCTTCTATAGAG 1320
Qy 1321 TTTTGGGACCGGCTCAGTGTGTTGAAGACACGAGCCGGCTTTTGTCAAGCGCACAC 1380
Db 1321 TTTTGGGACCGGCTCAGTGTGTTGAAGACACGAGCCGGCTTTTGTCAAGCGCACAC 1380
Qy 1381 GCTGATGAGTCCGCTTGTGTGTGAGAGCACTGTTGTTGTTGTTGTTGTTGTTGTTG 1440
Db 1381 GCTGATGAGTCCGCTTGTGTGTGAGAGCACTGTTGTTGTTGTTGTTGTTGTTGTTG 1440

Qy 1441 GAAGAGCCACGAGAGGAGAGAGTAACTAGACCCGAGAGATGATGAAATTACTGGGCTAC 1500
Db 1441 GAAGAGCCACGAGAGGAGAGAGTAACTAGACCCGAGAGATGATGAAATTACTGGGCTAC 1500
Qy 1501 TTTTGTCCGAACCGGAGATCTTATGAGAACGACCTGTCTGTGTGCGCAGCTTATATCTG 1560
Db 1501 TTTTGTCCGAACCGGAGATCTTATGAGAACGACCTGTCTGTGTGCGCAGCTTATATCTG 1560
Qy 1561 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCCCTCGACAGAGACTGAAAGAACCG 1620
Db 1561 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCCCTCGACAGAGACTGAAAGAACCG 1620
Qy 1621 CGGGTGAATTTTGAACAGACACATCCCGCTGATCTGTGTGCTCCGACATGCTCCAC 1680
Db 1621 CGGGTGAATTTTGAACAGACACATCCCGCTGATCTGTGTGCTCCGACATGCTCCAC 1680
Qy 1681 AGTCTCTTCTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1681 AGTCTCTTCTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 1741 CCTTGA 1746
Db 1741 CCTTGA 1746

RESULT 2
US-10-023-515-1
; Sequence 1, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; INVENTOR: Siles-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-023-515-1

Query Match 100.0%; Score 1746; DB 4; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCACAGGAGCTTATCATCTGCTTCAACATGATGTTTCTGATTTCTCCAGCCC 60
Db 96 ATGCCACAGGAGCTTATCATCTGCTTCAACATGATGTTTCTGATTTCTCCAGCCC 155
Qy 61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAGAGGCGCAGAGAGAAC 120
Db 156 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAGAGGCGCAGAGAGAAC 215
Qy 121 ACCAGGCTGGGATGATTCAGAGGCAAGACATCTGTGTGGAAAGCCCTGTGCTGTG 180
Db 216 ACCAGGCTGGGATGATTCAGAGGCAAGACATCTGTGTGGAAAGCCCTGTGCTGTG 275
Qy 181 AACGTGTTCTTCTGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGAGATTATGCAAC 240
Db 276 AACGTGTTCTTCTGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGAGATTATGCAAC 335

241 CCGACGCTGACGCGCCCTGGAGTAATTGCGAGAGACCACTCTACCTTAATTTGTGC 300
Db CCGACGCTGACGCGCCCTGGAGTAATTGCGAGAGACCACTCTACCTTAATTTGTGC 395
Qy 301 CTCGAGAACTCGAGGTGGCTGCTTTAGATCAACATGCTCAAGGTGATTAACCGAAA 360
Db CTCGAGAACTCGAGGTGGCTGCTTTAGATCAACATGCTCAAGGTGATTAACCGAAA 455
Qy 361 TTCCGAGGTGAGAGACTGGCTCTACCTGAACATCTAAGGCGCCGACCGCAATACA 420
Db TTCCGAGGTGAGAGACTGGCTCTACCTGAACATCTAAGGCGCCGACCGCAATACA 515
Qy 421 GAGCTCAAGCTCCCGCTGTGGTGTGTCCGAGAGGTGCTTGAAGAAGTGGCTCAGCC 480
Db GAGCTCAAGCTCCCGCTGTGGTGTGTCCGAGAGGTGCTTGAAGAAGTGGCTCAGCC 575
Qy 481 TCCATCTTTGATGGGTCCGCTGCTGCTGCTAATGAGAGCTGCTGGTGTGTGCTCAG 540
Db TCCATCTTTGATGGGTCCGCTGCTGCTGCTAATGAGAGCTGCTGGTGTGTGCTCAG 635
Qy 541 TACCGGCTAGAGAAATTTGGTTTCTTCAACACATGGGATCAGCATGCTCCGGGAACTGG 600
Db TACCGGCTAGAGAAATTTGGTTTCTTCAACACATGGGATCAGCATGCTCCGGGAACTGG 695
Qy 601 GCCTTCAAGACCAAGGTGGCTGCTGTCTGGGTCCAGAAAGACATCGAGTTCTTCGGT 660
Db GCCTTCAAGACCAAGGTGGCTGCTGTCTGGGTCCAGAAAGACATCGAGTTCTTCGGT 755
Qy 661 GGGGACCCGAGCTGTGAGACATCTTTGGGAGTCCGCGGAGGCCATAAGTGTTCAGT 720
Db GGGGACCCGAGCTGTGAGACATCTTTGGGAGTCCGCGGAGGCCATAAGTGTTCAGT 815
Qy 721 CTTAATCTGTCTCCCATGAGCCCAAGGCTTAATCCCAAAAGCATCATGAGAGTGGGGTG 780
Db CTTAATCTGTCTCCCATGAGCCCAAGGCTTAATCCCAAAAGCATCATGAGAGTGGGGTG 875
Qy 781 GCATCATCTCCCTTACCTGAGGAGCCATGATTAAGAGAGAGAGGAGCTGAGAGGTGT 840
Db GCATCATCTCCCTTACCTGAGGAGCCATGATTAAGAGAGAGAGGAGCTGAGAGGTGT 935
Qy 841 GCACATTTCTGTGTAAACATGCGTCAAGCTGAGAGCCCTGAGAGGTGCTGAGAGACA 900
Db GCACATTTCTGTGTAAACATGCGTCAAGCTGAGAGCCCTGAGAGGTGCTGAGAGACA 995
Qy 901 AAACCTCTCAAGAGCTGTGACCTCAAGCCCAAGAAACAAAGTCTTCACTGAAGTGT 960
Db AAACCTCTCAAGAGCTGTGACCTCAAGCCCAAGAAACAAAGTCTTCACTGAAGTGT 1055
Qy 961 GATGGTGTCTTCTTCTAATGAGCCTGATGATCTATTGCTCAGAAAGCATTTAAAGCA 1020
Db GATGGTGTCTTCTTCTAATGAGCCTGATGATCTATTGCTCAGAAAGCATTTAAAGCA 1115
Qy 1021 ATTCTTCATCATCGAGATCAATAACACAGAGTGGCTTCTGCTGCTAATGAAGAG 1080
Db ATTCTTCATCATCGAGATCAATAACACAGAGTGGCTTCTGCTGCTAATGAAGAG 1175
Qy 1116 ATTCTTCATCATCGAGATCAATAACACAGAGTGGCTTCTGCTGCTAATGAAGAG 1175
Db ATTCTTCATCATCGAGATCAATAACACAGAGTGGCTTCTGCTGCTAATGAAGAG 1260
Qy 1081 GCTCTGAGATCTCTAGTGGCTTCAACAGAGTCCCTTCCCTCATCTGATTAACAAACATC 1140
Db GCTCTGAGATCTCTAGTGGCTTCAACAGAGTCCCTTCCCTCATCTGATTAACAAACATC 1235
Qy 1176 GCTCTGAGATCTCTAGTGGCTTCAACAGAGTCCCTTCCCTCATCTGATTAACAAACATC 1235
Db GCTCTGAGATCTCTAGTGGCTTCAACAGAGTCCCTTCCCTCATCTGATTAACAAACATC 1320
Qy 1201 TCCCTGACTGAATTCGAGACAGTCTTCTGAGCTTCTGAGAGTGTCTTCTTGTGTGTC 1260
Db TCCCTGACTGAATTCGAGACAGTCTTCTGAGCTTCTGAGAGTGTCTTCTTGTGTGTC 1355
Qy 1261 CCTGACATGATCAAGCTGATTAACAGAGATGCTGTGAGACCTGTCTAATCTTAATGAG 1320
Db CCTGACATGATCAAGCTGATTAACAGAGATGCTGTGAGACCTGTCTAATCTTAATGAG 1415
Qy 1356 CCTGACATGATCAAGCTGATTAACAGAGATGCTGTGAGACCTGTCTAATCTTAATGAG 1415
Db CCTGACATGATCAAGCTGATTAACAGAGATGCTGTGAGACCTGTCTAATCTTAATGAG 1510
Qy 1321 TTTGGGACACGGGCTCAGTGTCTTGAAGACAGAGACCGGCTTTTGTCAAAGCCGACAC 1380
Db TTTGGGACACGGGCTCAGTGTCTTGAAGACAGAGACCGGCTTTTGTCAAAGCCGACAC 1475

1416 TTTGGGACACGGGCTCAGTGTCTTGAAGACAGAGACCGGCTTTTGTCAAAGCCGACAC 1475
Qy 1381 GCTGATGAAGTCCGCTTTGTGTTCGGTGTGCTTCTGTAAGGGGAGCATTTTATGTTTC 1440
Db GCTGATGAAGTCCGCTTTGTGTTCGGTGTGCTTCTGTAAGGGGAGCATTTTATGTTTC 1535
Qy 1441 GAAGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATGAATAATGAGGTAC 1500
Db GAAGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATGAATAATGAGGTAC 1595
Qy 1501 TTTGCTGAAACCGGGAATCTTAATGGAAACGACTGTCTGTGTGCTCAGCTTAATCTG 1560
Db TTTGCTGAAACCGGGAATCTTAATGGAAACGACTGTCTGTGTGCTCAGCTTAATCTG 1655
Qy 1561 ACTGAGCAGTACTCTCAGCTGATCTTGAACATGAGCTTGGACAGAGATCAAGAAACCG 1620
Db ACTGAGCAGTACTCTCAGCTGATCTTGAACATGAGCTTGGACAGAGATCAAGAAACCG 1715
Qy 1621 CGGATGATTTTGGACGACACATCCCGCTGATCTGTGCTCCGACATGCTCAC 1680
Db CGGATGATTTTGGACGACACATCCCGCTGATCTGTGCTCCGACATGCTCAC 1775
Qy 1681 AGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCAGACCTTTCTTTTCTTTGTGCT 1740
Db AGTCTCTTTCTTCTTAATCTTCTCTCTCTCTCAGACCTTTCTTTTCTTTGTGCT 1835
Qy 1741 CTTTGA 1746
Db CTTTGA 1841

RESULT 3
US-09-799-451-155
; Sequence 155, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yuning
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunhui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pl_FL_genes Version 2.0
; SEQ ID NO: 155
; LENGTH: 965
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(925)
; US-09-799-451-155

Query Match 29.1%; Score 508.6; DB 4; Length 965;
Best Local Similarity 81.0%; Pred. No. 3.2e-150;

Matches 675; Conservative 0; Mismatches 4; Indels 154; Gaps 1;

QY 91 GGGGCTTCTGCTGAAGGGCCACAGAGAACCCAGGCTGGAGTGGATTTCAGGGCAACAA 150
 Db 287 GGGGCTTCTGCTGAAGGGCCACAGAGAACCCAGGCTGGAGTGGATTTCAGGGCAACAA 346
 QY 151 GTCACTGTGCTGGAGAACCCCTGTGCTGTGACATGTTTCTCGAGTCCCTTTGCTGCT 210
 Db 347 GTCACTGTGCTGGAGAACCCCTGTGCTGTGACATGTTTCTCGAGTCCCTTTGCTGCT 406
 QY 211 CCCCCGTGGAGATCCCTGCGATTATTCAGAACCCCGAGCTGCTGCTGCGGATTAATTG 270
 Db 407 CCCCCGTGGAGATCCCTGCGATTATTCAGAACCCCGAGCTGCTGCTGCGGATTAATTG 466
 QY 271 CGAAGAGCCACCTCTACCCCTTAATTTGTGCTCAAGAACTCAAGATGCTGCTTAAAT 330
 Db 467 CGAAGAGCCACCTCTACCCCTTAATTTGTGCTCAAGAACTCAAGATGCTGCTTAAAT 526
 QY 331 CAACACATGCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAGAACTGCTCTACCTG 390
 Db 527 CAACACATGCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAGAACTGCTCTACCTG 586
 QY 391 AACATATATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
 Db 587 AACATATATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
 QY 451 CCAGAGAGTGTGCTTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
 Db 647 CCAGAGAGTGTGCTTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
 QY 511 TATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
 Db 707 TATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
 QY 571 ACATGGAGTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
 Db 766 ----- 765
 QY 631 TGGGTCCAGAGAGATGATGTTCTTGGTGGAGACCCAGCTGTGACATCTTTGGC 690
 Db 766 ----- 765
 QY 691 GAGTCCGGGGAGCCATAAGTCTTTCTAGTCTTAATCTGCTCTCCATGCGCAAGGCTTA 750
 Db 766 ----- -ATACTGTCTCCATGCGCAAGGCTTA 792
 QY 751 TTCCCAAGGCTCATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 810
 Db 793 TTCCCAAGGCTCATGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 852
 QY 811 TATGAGAGTGTGAGAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
 Db 853 TATGAGAGTGTGAGAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
 QY 871 TCTGAGGCTGCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
 Db 913 TCTGAGGCTGCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 965

RESULT 4
 US-09-949-016-3799
 ; Sequence 3799, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CU001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 3799
 ; LENGTH: 2117
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-3799

Query Match 23.3%; Score 406.6; DB 4; Length 2117;
 Best Local Similarity 56.5%; Pred. 1.2e-117;
 Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

QY 116 GGACACACAGGCTGGATGATGATTCAGGGCAAGCAAGTCACTGTGCGGAAGCCCTGTC 175
 Db 157 GGACACACACAGGCTGGATGATGATTCAGGGCAAGCAAGTCACTGTGCGGAAGCCCTGTC 216
 QY 176 CTGTGAAGTGTCTCTGGAATCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235
 Db 217 GGGTCCAAACCTTCTGGGAATTCATTTGCCAAGCCACTCTAGTCCGCTGGATTG 276
 QY 236 CGAACCCGAGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
 Db 277 CACCCCTGAGCCCTCTGATCTTGAATGTGAGAGTGAAGAGCAACCCATCCGCCA 336
 QY 296 TGTGCTTCAAGACTCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 352
 Db 337 TGTGCTTCAAGACTCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
 QY 353 ACCGGAATTCGAGTGTCAAGAACTGCTCTTCACTGAAACATTAATGCTGCTGCTGCTGCTGCT 412
 Db 397 TCCCTTCGACTCAGTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456
 QY 413 CCGATACAGCTCCAAAGCTCCCGCTTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472
 Db 457 GCATGAAAGCTCTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
 QY 473 GCTAGCTCTCACTTTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
 Db 517 GATGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
 QY 533 TCGTCCAGTACCGGCTAAGAAATTTGCTTCTTCAACATGAGATGAGATGAGATGAGATGAGAT 592
 Db 577 TCACTCAAGTACCGGCTAAGAAATTTGCTTCTTCAACATGAGATGAGATGAGATGAGATGAGAT 636
 QY 593 GGAACCTGGGCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652
 Db 637 GGAACCTGGGCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
 QY 653 TCTTGGTGGAGACCCAGCTGTGTAACAATCTTTGGCAAGTCCGCGGAGCCATTAAGT 712
 Db 697 ACTTTGAAGCAACCTGACCGGTGCAATTTTGGCAAGTCCGCGGAGCCATTAAGT 756
 QY 713 TTTCTAGTCTTAATGCTCTCCAGGCTTAATTCACCAAGCAATCAATGAGAA 772
 Db 757 TGTCTTGTGCTGT 816
 QY 773 GTGGGCTGCTCATGCTCTTCACTGAGAGCCCATGATTATTAAGAGTGAAGACCTTGC 832
 Db 817 GTGGGCTGCTCATGCTCTTCACTGAGAGCCCATGATTATTAAGAGTGAAGACCTTGC 870
 QY 833 AGGTGTGCTCAATTTCTGTGTGAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
 Db 871 CGGTGTGCTCAACCTGTGTGTGAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
 QY 893 TGAGCAAAACCTTCAAGAGCTGTGAACCTGAGCAAGAAACAAAGTCTTCACTGC 952
 Db 931 TGAGGAGCAAGATTAAGAGATTTCTGCAATTAACAAGCTTTCAAGATGATCCCG 990
 QY 953 GAGTGTGATGCTTCTTCTTAATGAGCTCTAAGATCTAATGCTCTCAAGAAAGAT 1012

Db 991 GAGTGTGATGGGGCTTCTCTGCCAGGCAACCCAGAGCTGTGGCTTGGCACT 1050
Qy 1013 TTAAGCAATTCCTTCATCATCGAGTCAATTAACGAGTGTGCTTCGCTGCCA 1072
Db 1051 TTCACCTGTCTCCAGCATGTGTGGTCAACAACAATGAATTCGCTGGCTCACTCCA 1110
Qy 1073 TGA-----AGAGGCTCTGAGATCTCGAGTGTGCTCAACAAGTCCCTTGGCC 1120
Db 1111 AGGTGATGAGATCTATATACCCAGAGAAATGAGACAGAGAGGCTCCAGGCTGCTC 1170
Qy 1121 TCCATCTGATACAAAACATCTGCAACATCCGCTCAAGTATTGCACTTTGGCTATG 1180
Db 1171 TGCAGAAATGTTAAGCTGTGATGTGCTCTCAATTTGTGTGACCTGTGAGGGAGG 1230
Qy 1181 AATACTTCATGACAGAGCACTCCCTGACCTGAATTCGAGACAGTCTTCTGACCTTGG 1240
Db 1231 AGTACATTTGGGAGCAATGGGGATCCCGACAGCCCTCCAGCGCATTCACAGAGATGATG 1290
Qy 1241 GAGATGTCTTTTGTGGTCCCTGCACTGATCAAGCTGATATGACAGAGATGTGTG 1300
Db 1291 CGGATCCATGTGTGTGATCCCTGCACTCCAGTACATTTTCAAGT---TTCGGGG 1347
Qy 1301 CACCTGTCTACTTCTATGATGTTTGGCACCGGCTCACTGCTTTTGAAGACAGAAAGCGG 1360
Db 1348 CCCCTGTACTTCTACGAGTTCAGACATCAGCCAGCTGGCTCAAGAAATCAGAGCCAC 1407
Qy 1361 CTTTGTCAAGGCGACACGCTGATGAGTCCGCTTGTGTGCTGTGTGCTTCTTGA 1420
Db 1408 CGCAGATGAGGCAACCAATGATGAGCTTCTTTGTTTCAAGAACTTTCTTTGGGG 1467
Qy 1421 AGGGGCAATTTGTTATGTTTCAAGAGACCAAGAGAGAGAAATTAAGTACAGCCGGA 1480
Db 1468 GCACTACATTTAAT-----CACTGAGAGAGAGAGAGAGCTTAAGCAGAGAA 1515
Qy 1481 TGATGAAATCTGGGCTACCTTTGCTGAAACCGGAATCTTAATGAGAAAGCACTGTCTC 1540
Db 1516 TGATGAAATCTGGGCTACCTTTGCTGAAATGAGAAACCCCAATGGCGAGGTCTGAC 1575
Qy 1541 TGTGGCCAGCTTATATCTGACTGAGCACTCACTGAGTGAATCTTGAACATGAGACTG 1600
Db 1576 ACTGCTGCTGTTCAGCAGAGAGAGCAATCTGAGTGAACCTTACAGCTGCGGTG 1635
Qy 1601 GACAGAGACTCAAGAACCGCGGTGATTTTGTGACCGACCATCCCC 1651
Db 1636 GCCGGCTCTGAGAGCCACAGGCTTCAGTTCTGAGAAAGGCGCTGCC 1686

RESULT 5
US-09-949-016-555
; Sequence 555, Application US/09949016
; Patent No. 6812319
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 555
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-555

Query Match 23.3%; Score 406.6; DB 4; Length 2169;

Best Local Similarity 56.5%; Pred. No. 1,2e-117;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;
Qy 116 GGAACACCAAGGCTGGATGATTCAGGCGCAGCAAGTCACTGTGCTGGGAACCCCTGTGC 175
Db 157 GGAACACCAACAGGCGGAGGTGTGGGAGCTTGTTCATGTGAAGGGCGCAATGCGG 216
Qy 176 CTGTGAACGTGTTCTCTGAGTCCCTTGTGCTCTCCCGCTGGAGATCCCTGCATTTA 235
Db 217 GGGTCCAAACCTTCTGGGAATTCATTTGCCAACCACTTATAGTCCGCTGCATTTG 276
Qy 236 CGAACCGGACCTGCATGCGCTGGATTAACCTTGCAGAAAGCACTTCTTACCTTAAT 295
Db 277 CACCCCTGAGCCCTCTGAATCTTGAATGTGTGAGGGAATGAACCAACCATCCGCCCA 336
Qy 286 TGTGCTTCAAGAACTCAAGTGTGCTTTGA---TCAACATATCTCAAGTGCAT 352
Db 337 TGTGTCTAAGGACCTCAACCGAGTGAAGTCAAGTCTTGAAGCAATTAAGTCACT 396
Qy 353 ACCGAAATTCGAGGTCTAGAAAGTGCCTTACCTGAACATCTATGCGCTGCCACG 412
Db 397 TCCCTTCCGATCTCATGTCTAGAGACTGCTGTACTTCAAGATCTTACAGCGCGCCATA 456
Qy 413 CGGATACAGGCTCAAGCTCCCGCTTGTGTGTGTTCCAGAGAGTGTCTTCAAGACTG 472
Db 457 GCCATGAAGGCTTCAACCTGCGGTGATGTGTGATCAAGGTGTGTGCTTGTGTTG 516
Qy 473 GCTCAGCTCCATCTTGTATGAGTCCGCTGTGCTGCTTATAGAGACGTGTGTGTG 532
Db 517 GCATGGCTTCTTGTATGATGATTCATGTGTGCTTGTGGAACGTGTGTGTGATCA 576
Qy 533 TCGTCAGTACCGGCTAGAAATTTGTTTCTTCAACCAATGGATCAGATGCTCGG 592
Db 577 TCACTCAATACGCTGTGTGTCTGTGCTTCTTCAAGCATGGAACAGACGCAACG 636
Qy 593 GGAACCTGGGCTTCAAGACCAAGTGTGTCTGTCTCTGTGGTCCAGAAACATGAGT 652
Db 637 GCACTGGGGCTACCTGGACCAAGTGTGTGCTCAAGCTGCGGTCCAGAGATATGCGCC 696
Qy 653 TCTTGTGTGGGAGACCCAGCTGTGACATCTTTGGGAGTCCCGGGAGCATTAATG 712
Db 697 ACTTTGAGGCAACCTTGAACCGTGTCAACATTTTGGGAGTCTCGGGTGGCAGAGTG 756
Qy 713 TTTTCACTTATACGTGTCTCCATGAGCAAGGCTTATTCACAAAGCATCATGAGA 772
Db 757 TGTCTTGT 816
Qy 773 GTGGGATGGCATCATCTCTTACCTGAGGCCCATGATTAAGAAAGTGAAGACCTGC 832
Db 817 GTGGGATGGCATCTCTGACCGGCTCATTTGCAAGCTCAGCTGA-----TGCATCTCA 870
Qy 833 AGGTGTTCACATTTCTGTGTGAACAATGCTCAGACTTGAAGCCCTGTGAGAGTGC 892
Db 871 CGGTGTGTGCAACCTGT 930
Qy 893 TGAGGCAAAACCTTCAAGAGGCTGACCTCGAGCAGAAACAAAGTCTTCACTC 952
Db 931 TCGGGGCAAGATGAAGAGAGATTTTGAATTAAGGCTTTCAAGATGATTCGG 990
Qy 953 GAGTGTGATGAGTCTTCTTCTTCTTAATGAGCTCTGATCTATGTCTCAGAAAGAT 1012
Db 991 GAGTGTGATGAGGCTTCTTCTGCGCCAGGACCCCAAGAGTGTGGCTGTGCGACT 1050
Qy 1013 TTAAGCAATTCCTTCATCATCGAGTCAATTAACGAGAGTGTGCTTCTGCTGCTA 1072
Db 1051 TTCACCTGTCTCCAGCATGTGTGTGTCACAAACAATGAATTCGCTGGCTCACTCCA 1110
Qy 1073 TGA-----AGAGGCTCTGAGATCTCGAGTGTGCTCAACAAGTCCCTTGGCC 1120
Db 1111 AGGTGATGAGATCTATATACCCAGAGAAATGAGACAGAGAGGCTCCAGGCTGCTC 1170
Qy 1121 TCCATCTGATACAAAACATCTGCAACATCCGCTCAAGTATTGCACTTTGTGCTAATG 1180

Db	1171	TGCAGAAAATGTTAAAGCGTGTGATGATGTTGCCCTTCACATTTGGTATACCTGCTGAGGAGG	1230
Qy	1181	AATACTTCATGACAGACAGCTCCCTGACTGAAATCCGAGACAGTCTTGTGACCTTGTG	1240
Db	1211	AGTACATTGGGGACAAATGGGGATCCCGACAGCCCTCCAAAGCCGACAGTTCCAGAGATGATGG	1290
Qy	1241	GAGATGTGTTCTTTTGTGGTCCCTGACATGATACAGCTCCGATATACAGAGATGCTGTG	1300
Db	1291	CGAGCTCATGTTTGTGTGATTCCTTGACCTCCAAATGCAATTTTCAAGT---TTCGGGG	1347
Qy	1301	CACCTGTCTACTTCTATGATGTTTGGGACCGGGCTCACTGTCTTGAAGACAGAAAGCGG	1360
Db	1348	CCCCGTGTACTTCTACAGAGTTCACGACATCAGCCAGCTGGCTCAAGAAATCAGGCGAC	1407
Qy	1361	CTTTTGTCAAGCCGAGCCAGCTGATGAAAGTCCGTTGTGTTGGTGTGCTTCTCTGA	1420
Db	1408	CGCAGTAAAGGCGAACCATGTGATGAGCTTCTTTGTTTCAAGATTTCTTTGGG	1467
Qy	1421	AGGGGGACATTGTTATGTTTCGAAGAGGCCAGAGAGAGAGATTACTGAGCCGGAAGA	1480
Db	1468	GCAACTACATTAAATY-----CACTGAGGAAGAGACAGCTAAGCAGGAAGA	1515
Qy	1481	TGATGAATAATCTGGGCTACTTTTGTCTGAAACCGGGAATCCTAATGGGAAGACCTGTCTC	1540
Db	1516	TGATGAAGTATGGGGCCAACTTTTCGAGAAATGGGAAACCCCAAGGGGAGGGTCTGCGAC	1575
Qy	1541	TGTGGCCAGCTTATATCTGACTGAGAGATGACTCCAGCTGGAATTGAACATGAGACTCG	1600
Db	1576	ACTGGCCGCTGTTGCAACAGGAGGAAATACCTGACGTGAACCTAACAGCTCGCGTGG	1635
Qy	1601	GACAGAGACTCAAGAAACCGCGGTGGAATTTTGGACAGACCAATCTCCCC	1651
Db	1636	GCGGGCTCTGAAGGCCCAACAGGCTTCAGTTCTGGAGAAAGGCGCTGCCCC	1686

```

RESULT 6
US-09-595-682B-27.
; Sequence 27, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Poetter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-595-682B-27

```

Query Match	23.3%	Score 406.6;	DB 4;	Length 2191;
Best Local Similarity	56.5%;	Pred. No. 1.2e-117;		
Matches 876;	Conservative	0;	Mismatches 639;	Indels 36; Gaps 5;

QY	116	GGAA	CAC	CAGG	CTGG	AT	TG	AT	TGA	TTC	AGG	C	AG	CA	G	CA	AG	T	CA	T	CT	GT	GT	G	AA	AA	CC	CT	GT	C	175
Db	179	GGAC <td>CA <td>CAC <td>AC <td>CGGG <td>GA <td>GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>CAC <td>AC <td>CGGG <td>GA <td>GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CAC <td>AC <td>CGGG <td>GA <td>GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	AC <td>CGGG <td>GA <td>GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CGGG <td>GA <td>GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GGT <td>GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GT <td>CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CGGG <td>GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td></td>	TT <td>GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td></td>	GT <td>CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td></td>	CA <td>T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td></td>	T <td>AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td></td>	AT <td>T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td></td>	T <td>GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td></td>	GA <td>AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td></td>	AGG <td>CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td></td>	CC <td>CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td></td>	CA <td>T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td> </td>	T <td>CCG</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>238</td>	CCG							238
QY	176	CTG <td>GA <td>CA <td>CG <td>AT <td>T <td>CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>CA <td>CG <td>AT <td>T <td>CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>CG <td>AT <td>T <td>CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CG <td>AT <td>T <td>CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	AT <td>T <td>CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	T <td>CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CG <td>AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	AG <td>T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	T <td>CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	T <td>T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	T <td>GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GT <td>GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	GT <td>CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td></td>	CC <td>CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td></td>	CC <td>CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td></td>	CC <td>CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td></td>	CT <td>GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td></td>	GG <td>AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td></td>	AT <td>CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td></td>	CC <td>CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td></td>	CT <td>GG <td>AT <td>TT <td>235</td> </td></td></td>	GG <td>AT <td>TT <td>235</td> </td></td>	AT <td>TT <td>235</td> </td>	TT <td>235</td>	235
Db	239	GGGT <td>CA <td>AA <td>CC <td>TT <td>CT <td>CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>AA <td>CC <td>TT <td>CT <td>CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	AA <td>CC <td>TT <td>CT <td>CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>TT <td>CT <td>CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TT <td>CT <td>CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CG <td>GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TT <td>CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TT <td>TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	TG <td>CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CA <td>AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td></td>	AG <td>CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td></td>	CA <td>CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td></td>	CC <td>CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td></td>	CT <td>TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td></td>	TA <td>GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td></td>	GT <td>CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td></td>	CC <td>GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td></td>	GT <td>CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td></td>	CG <td>CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td></td>	CA <td>TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td> </td>	TT <td>TG</td> <td></td> <td></td> <td></td> <td>298</td>	TG				298	
QY	236	CGAA <td>CC <td>GA <td>CC <td>CT <td>GA</td> <td>TC <td>CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>GA <td>CC <td>CT <td>GA</td> <td>TC <td>CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GA <td>CC <td>CT <td>GA</td> <td>TC <td>CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>CT <td>GA</td> <td>TC <td>CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>GA</td> <td>TC <td>CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	GA	TC <td>CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CG <td>CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td></td>	CC <td>CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td></td>	CT <td>GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td></td>	GG <td>GA</td> <td>TA</td> <td>CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td></td>	GA	TA	CT <td>TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td></td>	TG <td>CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td></td>	CA <td>AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td></td>	AA <td>GA</td> <td>CA</td> <td>CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td></td>	GA	CA	CC <td>CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td></td>	CT <td>CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td></td>	CT <td>TA</td> <td>CC</td> <td>CT <td>TA</td> <td>TT</td> <td></td> <td>295</td> </td>	TA	CC	CT <td>TA</td> <td>TT</td> <td></td> <td>295</td>	TA	TT		295	

[illegible]

Db 1370 CCCCTGTGACTTCTACGAGTTCAGCATCAGCCAGCTGCTCAAGAACATCAGGCCAC 1429
Qy 1361 CTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTTGTGTGGTGTGCTTCTCA 1420
Db 1430 CGCATGAAAGGACAGACATGATGATGACCTTCTTTTTCAGAACTTTCTTTGGGG 1489
Qy 1421 AGGGGACATTTGTTATGTTTCAGAAAGCCAGGAGAGAGAAATTACTAGCCGAGAA 1480
Db 1490 GCAACTACATTAAAT-----CACTGAGGAAGAGAGAGAGCTAAAGCAGAGAA 1537
Qy 1481 TGATGAATATCTGGGCTACCTTTGCTCGAACCGGAAATCTTAATGGAACGACTGTCTC 1540
Db 1538 TGATGAATATCTGGGCTACCTTTGCTCGAACCGGAAATGGAACCGGAGGCTGTGCCAC 1597
Qy 1541 TGTGGCCAGCTTATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1600
Db 1598 ACTGCGCGCTGTTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1657
Qy 1601 GACAGAGACTCAAGAAAGCCGAGTGTGATTTTTCAGACAGACATCCCCC 1651
Db 1658 GCGGCGCTCTGAAGGCCACAGAGCTTCAGTTCTGGAAGAGCGCTGCCCC 1708

RESULT 7
US-09-595-682B-20
; Sequence 20, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Poter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; Tumor Cells
; FILE REFERENCE: SU-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: *Oryctolagus cuniculus*
US-09-595-682B-20

Query Match 20.8%; Score 362.6; DB 4; Length 1717;
Best Local Similarity 54.8%; Pred. No. 8,8e-104;
Matches 880; Conservative 0; Mismatches 649; Indels 78; Gaps 5;
Qy 91 GGGCCTTGTGCTGAAGGGCCAGAGAACACAGAGCTGGATGATTCAGGGCAAGCA 150
Db 62 GGGGACCGGTGACCACTGTGTGATGATGATGATGATGATGATGATGATGATGATG 121
Qy 151 GTCACTGTGCTGGAGAGCCCTGTGCTGAACTGTTCTTGGAGTCCCTTTGCTGCT 210
Db 122 GTCACTGTGCTGGAGAGCCCTGTGCTGAACTGTTCTTGGAGTCCCTTTGCTGCT 181
Qy 211 CCCCCGTGGGATCCCTGATTTTGAACCCGAGCGCTGATGCGCCCTGGGATTAATTTG 270
Db 182 CCCCCGTGGGATCCCTGATTTTGAACCCGAGCGCTGATGCGCCCTGGGATTAATTTG 241
Qy 271 CGAGAGCCAGCTCTTACCTTAATTTGCTTCAAGAA-----TCAGAGTGGCTGCTC 324
Db 242 AAGAACACCACTCTTACCTTCCATGTGCTTCCAGAGAGAGAGAGAGAGAGAGAGAG 301
Qy 325 TTAGTCAACAGAGCTCAAGTGCATTAACCGAAATTCGAGAGTGCAGAGAGAGAGAG 384
Db 302 TCGAGCTCTTCAACCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Qy 385 TACTGAACATCTATGCGCTGCCACAGCGAGATACAGGCTCAAGCTCCCGTCTTGGTG 444

Db 362 TACTGAATATTTACACCCCTGCTGACCTGACAAAGAGAGAGAGAGAGAGAGAGAGAG 421
Qy 445 TGGTCCCAAG 504
Db 422 TGGATTCATGAG 481
Qy 505 GCTGCTTATGAG 564
Db 482 TGTGCCCATGAG 541
Qy 555 TTACCAACATGAGAGATGAG 624
Db 542 TTACCAACATGAG 601
Qy 625 CTGTCTGTGCTTCAAG 684
Db 602 CTGCGGTGGGTTCAG 661
Qy 685 TTTGCGAGAGTCCGCGGAG 744
Db 662 TTTGAGAGAGTCAAG 721
Qy 745 GGTATATTCACAAAGCCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804
Db 722 AATCTCTTCATGAG 781
Qy 805 CATGATTTATGAG 864
Db 782 AAGAACACCAAGTCTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
Qy 865 TCAGACTGAG 923
Db 842 TCGGCTGATGAGTTCATGCTGCTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
Qy 924 -----CTTCAGGACAG 945
Db 902 TTGAAGATGAATTTATGAGCTCTGATGATGATGATGATGATGATGATGATGATGATG 961
Qy 946 TTCACTCAAG 1005
Db 962 CTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1021
Qy 1006 AAGAGATTTAAAGCAATTCCTTCATCATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
Db 1022 AAGAAATACAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1081
Qy 1066 CTGCTATGAG 1114
Db 1082 ATCCCAATGCAAGATCTGAG 1141
Qy 1115 -TTGCCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1173
Db 1142 ACAGAACTCTGTGAG 1201
Qy 1174 GCTAATG---AATATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230
Db 1202 GCACTGAGAGAGATTTGAG 1261
Qy 1231 GACTTGTGAG 1290
Db 1262 GACATGCTTGGAG 1321
Qy 1291 GATGCTGTGAG 1350
Db 1322 GATGCTGTGAG 1381
Qy 1351 ACGAG 1410
Db 1382 ATGAG 1441
Qy 1411 GCTTCTGAG 1470

Db 1442 CCGTTTTTAA-----AGAGGTCCACAGAAAGAGATCAAACTG 1483
QY 1471 AGCCGGAAGATGATGAATACTGGCTACCTTGGCTCGAACCAGGAATCTTAATGGAAAC 1530
Db 1484 AGGAAGATGGTGAATGAATACTGGCCCACTTTGCTGAGATGGAAATCCCAATGAGAA 1543
QY 1531 GACCTGCTCTGTGGCCAGCTTAACTGATCTGAGAGATCCTCGAGCTGACTTGAAC 1590
Db 1544 GGGCTTCTCAATGGCCAGCATATGACTACAGAAAGTTACCTGCAAGATTGGAGCAC 1603
QY 1591 ATGAGCTGGAGAGACTCAAGAACCCGGGGTGGATTTTGGAC 1637
Db 1604 ACCAGGACCCAGAAATGAAAGACAAAGAAAGTGGCTTTTGGAC 1650

RESULT 8
US-09-264-737-3
; Sequence 3, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; FILE REFERENCE: 38-21(10551) RLB3 Pyridine Tolerance
; CURRENT APPLICATION NUMBER: US/09/264,737A
; EARLIER FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Rabbit
US-09-264-737-3

Query Match 20.4%; Score 357; DB 3; Length 1701;
Best Local Similarity 54.8%; Pred. No. 5.2e-102;
Matches 866; Conservative 0; Mismatches 635; Indels 78; Gaps 5;

QY 119 ACACGAGCTGGAGATGATTCAGGCAAGCAAGTCACTGTGCGGAGCCCTGCTG 178
Db 83 ACACGTGAAGGGAAGTCTCGGGAAGTTCGTGACTTGAAGAGATTTGCACAGCCG 142
QY 179 TGAAGTGTCTCTGGAGTCTCTTGTCTGCTCCCGCTGGAGTCCCTGGATTTA 238
Db 143 TGGCCGTCTCTGGAGTCTCTTGCCTTGCAGAGCCCTTGGATCTCTGAGTTG 202
QY 239 ACCGCGAGCTGCAATCGCCCTGGAGTAACTTGCAGAAAGCACTCTCTAATTTGT 298
Db 203 CACCAAGCTGCAAGATCAAGAGCCAGTGAAGAACACCTCTCACTCCCATGT 262
QY 299 GCTTCCAGAAC-----TCAGAGTGGCTGCTTGAATCAACATGCTCAAGGTG 352
Db 263 GCTCCCAAGAGAGATCAAGAGCAATGCTCTGAGAGCTTTCAACAAGAAAGAA 322
QY 353 ACCGGAATTTGGAGTGCAGAAAGTCTCTAATCTGAACATGTATGCTGCCAGC 412
Db 323 ACATCCCTCTTAAGTTTCTGAAGACTGCTTTACTGAATATTACACCTCTGACC 382
QY 413 CCGATACAGGCTCCAGAGTCCCGCTTGGTGTGGTTCCAGAGAGTCCCTTAAGACTG 472
Db 383 TGAACAAGAGAGAGAGCTGCGGTGATGTGTGATTCATGAGGTGTGTGATGG 442
QY 473 GCTTACGCTTCATTTTGAATGGGTCCGCTGCTGCTTGAAGAGTCTGGTTTGG 532
Db 443 GTGAGAGATCAACTATGATGGCTGCTCTTCTTGTGCCATGAAAGTGTGTGTGA 502
QY 533 TCGTCCAGTACCGGCTGGAATATTTGTTTTCACCAATGGATCAGACTGCTCCG 592
Db 503 CCAATTCATGATCCGCTGGGATCTGGGATTTCTACGACAGAGATGAGCACAGCCG 562

QY 593 GGAATCGGCTTCAAGACCAAGTGGCTCTGTCTGTGCTCGAAGAAATCATGAGT 652
Db 563 GGAATCGGCTTCAAGACCAAGTGGCTCTGTGCTCGAAGAAATCATGAGT 622
QY 653 TCTTGGTGGAGACCCAGAGCTGTGACATCTTTGGAGAGTCCGGAGACATTAATG 712
Db 623 ACTTTGAGAGGAGACCAAGAGCTGTGACATCTTTGGAGAGTCCGGAGACATTAATG 682
QY 713 TTTTGAATCTTAATCTGTCTCCATGAGCAAGGCTTATTCACAAAGCTCATGAGAG 772
Db 683 TCTTATCTCTTATTAATCTCCCTGACCAAGAAATCTCTCATGAGCAATTTCCGAGA 742
QY 773 GTGGAGTGGCATATCTCCCTTAACCTGAGGCCATGATTAAGAGAGTGAAGCTG 832
Db 743 GTGGAGTGGCATCTCTTTCAGTCTTTCAGAGAAAGAACCAAGTCTTGGCTGAGAAA 802
QY 833 AGGTGTGACATTTCTGTGTAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 892
Db 803 TTGCATGAAAGTGGGTAAACCAACACTGCGCTGTATGATGATGATGATGATGATG 862
QY 893 TGAAGACAAACCTCCAGAGAGTCTGAC----- 923
Db 863 AGAAGACAGAGAAAGAACTCATGAGGTGACATTAAGAAATTAATGCTGATGATG 922
QY 924 -----CCTCAGCCAGAAACAAAGCTTTCACCTCAGTGTGATGATGATGATG 973
Db 923 TAGTTGGAGACCCAAAGAAACACCGCTTCTGACCACTGATGATGATGATGATGATG 982
QY 974 TTTCTAATGAGCTCTGATCTAATTTGTCTGAGAAAGATTTAAAGCAATTCCTTCA 1033
Db 983 TGCCAAAGACCTGACAGATTTCTGACAGAGAAAGAAATCAATGCTCCTCACTAC 1042
QY 1034 TCGAGTCAATTAACCAAGATGTGGCTTCTGCTGCTTATGAAGAGCTCTGAGATTC 1093
Db 1043 TGGAAATCAACCAAGATTTGGCTGATTTCCCAATGCAATGCTGCTATTCAC 1102
QY 1094 TCAATGCTCCAAAGATCTCT-----GCCCTCAATGATCAAAATCAATC 1141
Db 1103 TCTCTGAAGCAACTGACCAAGACACTCAAGAACTTTGTGAGTCTTACCCCA 1162
QY 1142 TGCAATCCGCTCAGATTTTGAAGCTTGTGCTAATG---AATCTTCATGACAGC 1198
Db 1163 TTGTCAATGTCTTAAGAGAGTCACTCAAGTGCACATGAGATTTTGAAGAGAGC 1222
QY 1199 ACTCCCTGAAGAAATCCAGACAGTCTTCTGAGATTTGTGATGATGATGATG 1258
Db 1223 ATGACCTGTCAAAAGAAAGACTTTTCTTCTGAGATGCTTGAATTTGATTTG 1282
QY 1259 TCCCTGACATGATCAGAGTCTGATATCAAGAGATGCTGTGACCTGTACTTATG 1318
Db 1283 TCCCATCTGGAATGTGGCTGTCACCAAGAGATGCTGAGGCCCTCACTATATG 1342
QY 1319 AGTTTGGACCGGCTCAGTGTCTTGAAGACAGAGCCGCTTTTGAAGCCGAC 1378
Db 1343 AGTATGTATGCTCCAGACTTCTATCAATGAGACCCAAAGACAGATGAGGAGC 1402
QY 1379 ACCTGTGAAGTCCGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 1438
Db 1403 ATGAGATGAGATCTTCTGTCTTGAAGCCCTTTTAAAGAG----- 1449
QY 1439 TCGAAGAGCCACAGAGAGAGAAATTAAGAGCCAGAAAGTATGAAATCTGAGCT 1498
Db 1450 -----GTGCTCAAGAAAGAGATCAAACTGAGCAAGATGATGAAATCTGGGCA 1504
QY 1499 CTTTGTCTGAACCGGGAATCTTAATGGAAGAGACTGTCTTGTGGCACTTAATC 1558
Db 1505 ACTTTGTGAATGGAATGGAATCCCAATGAGAGGCTTCTCAATGAGCAGATATGCT 1564
QY 1559 TGAATGAGAGTACCTCCAGCTGAGCTTGAACATGAGCTCGAGCAGAGCTCAAGAG 1618
Db 1565 ACAAGGAAGATTAAGTGAAGATTTGAGAGCACCAGCAGCAGCAGAACTGAAGACA 1624
QY 1619 CGCGGATGATTTTGGAC 1637

Db 1625 AGGAAGTGGCTTCTGGAC 1643

RESULT 9

```

US-09-810-861B-5
: Sequence 5, Application US/09810861B
: Patent No. 6770799
: GENERAL INFORMATION:
: APPLICANT: Mori, Tsafir S.
: APPLICANT: Soreq, Hermona
: APPLICANT: Arntzen, Charles J.
: APPLICANT: Maason, Hugh S.
: TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
: TITLE OF INVENTION: TRANSGENIC PLANTS
: FILE REFERENCE: BtI-45
: CURRENT APPLICATION NUMBER: US/09/810,861B
: CURRENT FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/190,440
: PRIOR FILING DATE: 2000-03-17
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn Ver. 3.1
: SEQ ID NO 5
: LENGTH: 1725
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURES:
: OTHER INFORMATION: Description of Artificial Sequence: synthetic
: OTHER INFORMATION: human acetylcholinesterase gene optimized for
: OTHER INFORMATION: expression in plants
: US-09-810-861B-5

```

Query Match	8.0%;	Score 139.4;	DB 4;	Length 1725;
Best Local Similarity	55.1%;	Prod No 4	8e-33;	

Qy	175	CCTGTGAACGATGTTCTCTGGAGATCCCTTTGGCTGCTCCCCGCTGGAGATCCCTGCATTT	234
Db	175	CCTGTCTCTGCTTCTCTGGGATCCCTCTTTGGGAGCCACCATGGGACCCCGTCGCTTT	234
Qy	235	ACGAACCCGACGCTGCATCGCTCTGGAGTAATCTTGGAGAACCACTTCTACCTTAAT	294
Db	235	CTGCACCCGAGGCCAAGACGCTTGGTCAAGGGGTGGTAGCGCTAACCTTCCAGAGT	294
Qy	295	TTGTGCTCTCCAGAACTCAGAGTGGGTGCTCTTAGATCAACAATGCTCAAGGTGCATTAC	354
Db	295	GTCTGCTTACCAATATGTGGACAACCTTATCCAGGTTTGAAGGGCAACGGAATGTGAAC	354
Qy	355	CCGAATAATCGAGTGTCAAGAAAGTCCCTTACCTGAACATCTATCGGCTGCCAGGCC	414
Db	355	CCCAACCGTGAAGTACGAGGACGCTCGTATCTTAAGGTGGACACATATCCCCGG	414
Qy	415	GATACAGGCTCCAGACTCCCGCTCTTGTGTGTGTTCCAGAGAGTCCCTTCAAGACTGAC	474
Db	415	CCTACATCCCA-----CCCTGTCTCGTCTGGATTAATGGGGTGGCTTTACAGTGGG	471
Qy	475	TCAGCTC-----CATCTTGAATGGGTCCGCTCGCTGCTTATAGACGTGCTGGTT	528
Db	472	GCTTCCTCTTGAAGATGTACATGGCCGCTTCTTGTATACGGCCGAGAGCATGTGCTG	531
Qy	529	GTGTGCTGTCCAGTACCCGCTAGGAATATTGTGTTCTTCAAC--ACATGGGATCAAGAT	585
Db	532	GTGTTCATGAATACCCGGGTGGAGCTTTGTGCTTCTGGCCCTGTCCGGGAGCCGAGAG	591
Qy	586	GCTCCGGGGAACCTGGGCTTCAAGGACCAAGTGGTCTCTGTCTGGGTCCAGAAAGAAC	645
Db	592	GCCCCGGGCAATGTGGGTCTCTGGATCAAGAGCTGGCCCTGAGATGGGTGACAGAGAAC	651
Qy	646	ATCGAGTCTTTCGATGGGACCCCAAGCTGTGACCATCTTTGGCGAGTCCGGGGAGCC	705
Db	652	GTGGGAGCCTTGGGGGGTAGCCGACATCAGTAGCGCTGTTTGGGAGAGCGGGAGGCC	711
Qy	706	ATTAAGTGTTTCTAATCTTAATCTGTCTCCCATGGCCAAAGGCTTAATCCAGAAAGCATC	765

Db 712 GCTTCGGTGGGCATGCACTGCTGTCCCGCCAGCCGCGGCGCTGTTCACACAGGCGCGTG 771

QY 766 ATGAGAGTGGCGTGCCCA 784
|||
Db 772 CTGCAGAGCGGTGCCCA 790

RESULT 1C
US-07-732

Sequence 1, Application US/07732962A
Patent No. 5248604
GENERAL INFORMATION:
APPLICANT: Fischer, Mel
TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White, Esq.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/732,962A
FILING DATE: 19910722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UT
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
US-07-732-962A-1

Query Match	8.0%	Score 139.4	DB 1	Length 1845
Best Local Similarity	55.1%	Pred. No. 56-33		
Matches 341	Conservative 0	Mismatches 266	Indels 12	Gaps 3

Qy	235	ACGAAACCGCAGCCTGCGATCGCCCTTGAGATACCTTGAGAAAGCCACCTCTCTACCTTAAT	234
Db	235	CTGCGACACCGAGGCCCAAGCAGCCTTGATGACGGGGTGTGATGACGCTACCAACCTTCAAGAT	234
Qy	175	CCGTGTAAAGTGTCTCTCGAAGTCCCTCTTGTGTCTCCCGCTGGGATCCCTTCGATTT	234
Db	175	CTGTCTCTGCTTCTCTGGGATCCCCCTTTGGAGAGCAACCAATGGAAACCCCGTGGCTTT	234
Qy	295	TTGTGCTCTCAAACTCAGAGTGGCTCTTAAATCAACACATGCTCAAGTGCATTAC	354
Db	295	GTCGTGTACCAATATGTGGACACCTTATACCAAGGTTTGTGGGCAACCGAGATGTGAAAC	354
Qy	355	CCGAAATTTGGAGGTGTCAAGAAAGTGGCTTACCGGAACATCTATGCGGCTGGCCAGGCC	414
Db	355	CCCAACCGTAAAGCTAGACGAGGATCGCTGTACCTCAACGTTGTGACACATATCCCTCGG	414

QY	415	GATATACAGGCTCCAAAGCTCCCGCTTGATGTGTTCCAGAGGTGCTTCAAGACTGGC	474
Db	415	CCTACATGCCCA---CCCTGCTCTGTGTGATCATATGGGGTGGCTTTACAGTGGG	471
QY	475	TCAGCTC-----CATCTTTATGAGGTCCGCCCTGTGCTCATATGAGACGTGCTGGTT	528
Db	472	GCCTCCTCTTGAGACGTATACATATGCGCTTCTTGGTACAGGCCAGAGACATGTGCTG	533
QY	529	GTGTGTCGTCCAGTACCGGCTAGGAATATTTGGTTTCTTACG---ACATGGATCAGCAT	585
Db	532	GTGTCCATGAACTACCGGATGGGAGGCTTTGGCTTCTGTGCCCTGTCCGGGAGACCGAG	591
QY	586	GCTCCGGGGACCTGGGCTTTCAAAGACCAAGGTGCTGTCTGTCTGTGGGTCCAGAAAGAC	645
Db	582	GCCCCGGGCAATGTGGGTCTCTGTGATCAGAGGCTGGCCTTGATGTGGGTGCAGAGAAC	651
QY	646	ATCGAGTTCTTCGTGTGGGAGCCCAAGCTGTGTACATCTTTTGGCAAGTCCGGGAGCC	705
Db	652	GTGGCAGGCTTCGGGGGTGATCCCGACATCAGTACGCTGTTTGGGAGAGCGCGGAGCC	711
QY	706	ATTAAGTGTTCATGTCTTAACTGTCTGTCTCCATGAGCCAAAGCTTAATTCACAAAGCATC	765
Db	712	GCCTGTGGTGGCATCAGCTGTGTCTCCGCCAGCGGGGCTGTTCACAGAGGCCGTG	771
QY	766	ATGAGAGATGGGGTGCCCA	784
Db	772	CTGCAGAGCGGTGCCCCCA	790

```

RESULT 11
PCT-US92-06106-1
Sequence 1, Application PC/TUS9206106
GENERAL INFORMATION:
APPLICANT: Fischer, Melr
TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White, Esq.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06106
FILING DATE: 19920722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 33304-B-PCT/JPM/EA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
PCT-US92-06106-1

```

Query Match	8.0%;	Score 139.4;	DB 5;	Length 1845;
Best Local Similarity	55.1%;	Pred. No. 5e-33;		
Matches 341;	Conservative 0;	Mismatches 266;	Indels 12;	Gaps 3;

175 CCTGTGAACGTGTTCCTCCGGAAGTCCCTTTTGCTGCTCCCGCTGGAAATCCCTGGATTT 234

175 CCTGTGAACGTGTTCCTCCGGAAGTCCCTTTTGCTGCTCCCGCTGGAAATCCCTGGATTT 234

175 CCTGTGTCTGTGCTTTCTGTGGGCATCCCTTTTGCGAGGCCAATGAGAACCCCGTGGCTTT 234

235 ACGAACCCCGACGCTGTGATGCGCTTGGGATTAATTGGAGAACCACTTCTTAACCTTAAT 294

235 CTGCAACCCGAGGCCCAAGCAGCCTTGTGTGAGGGGTGTGAAGCTTAACAACCTTCCAGAGT 294

295 TTGTGCTCTCCAGAACTCAGAGTGGCTGTCTTTAGATCAACATCTCAAGGTGATTAC 354

295 GTCTGTCAACAAATATGTGGACAACCTTAATCCAGAGTTTGAAGGACCGAAGATGTGAAAC 354

355 CCGAAATTTGGGAATGTTCAGAAAGATGTCCTTACCTGAACATTAATGCGCTGCCAGGCC 414

355 CCGAACCGTAAAGCTGAGACGAGACGTGCTCACTTCAACGTGTGGAACAAATACCCCGG 414

415 GATACAGGCTCCAAAGCTCCCGCTTTGGTGTGTGATTTCCAGAGAGTGCCTTCAAGACTGC 474

415 CTAATACATCCCCA---CCCTGTCTGCTGTGATCTAATGGGGGTGGCTTCAAGTGG 474

475 TTAGCCCTC-----CATCTTTGATGTGGGTCCGCGCTGTGCTTAAGAGACGTGTGTT 528

472 GCTCTCTCTTGGACGTGTACGATGGCCGCTTCTTGTATCAAGCCGAGAGACGTGTGTT 531

529 GTGTGTGTCCAGTACCGGCTAGGAATATTTGGTTTCTTCAAC---ACATGTGAATCAGCAT 585

532 GTGTTCATGAACACCGGGGTGGAGCTTTGGCTTCTCGGCCCTGACCGGGAGCGCAGAG 591

586 GCTCCGGGGAACTGGGCTTTCAAAGGACCAAGTGTGTGCTGTCTCTGTGGTCCAGAGAAC 645

592 GCCCGGGGCAATGTGGGTCTCTCTGTATCAGAGAGCTGGCCCTGTCAATGTGGTCCAGAGAAC 651

646 ATGGAAGTTCTTGTGTGGGGAACCCAGCTCTGTGACCAATCTTTGGCAAGTCCGCGAGACC 705

652 GTGGCAGCCTTGGGGGTGACCCGACATCAGTGAAGCTGTTTGGGGAGAGCGCGGAGACC 711

706 ATTAAGTGTTCATAGTTTATTACTGTCTCCATGCGCAAGGCTTAATTCACAAAGCCATC 765

712 GCTCGGTGGGATGTGACCTGTGTGTCCCGGCCAGCGGGGCTGTTCACAGGGCCGTG 771

766 ATGGAATGTGGGTGACCA 784

772 CTGCAAGCGGTGCCCA 790

```

RESULT 12
US-09-949-016-1192
; Sequence 1192, Application US/09949016
; Patent No. 681339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,766
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1192
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1192

```



```

1 OPERATING SYSTEM: PC-DOS/MS-DOS
2 SOFTWARE: Patent Release #1.0, Version #1.30
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/08/318,826A
5 FILING DATE:
6 CLASSIFICATION: 514
7 ATTORNEY/AGENT INFORMATION:
8 NAME: Kohn, Kenneth I.
9 REGISTRATION NUMBER: 30,955
10 REFERENCE/DOCKET NUMBER: 2391.00001
11 TELECOMMUNICATION INFORMATION:
12 TELEPHONE: (248) 539-5050
13 TELEFAX: (248) 539-5055
14 INFORMATION FOR SEQ ID NO: 5:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 2256 base pairs
17 TYPE: nucleic acid
18 STRANDEDNESS: double
19 TOPOLOGY: linear
20 MOLECULE TYPE: CDNA to mRNA
21 HYPOTHEetical: NO
22 ANTI-SENSE: NO
23 ORIGINAL SOURCE:
24 ORGANISM: Homo sapiens
25 FEATURE:
26 OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
27 3, 4 and 6"
28 US-08-318-826A-5

```

Query Match	8.0%	Score	139.4	DB 2	Length	2256			
Best Local Similarity	55.1%	Pred. No.	5.8e-33						
Matches	341	Conservative	0	Mismatches	266	Indels	12	Gaps	3

QY	175	CGTGTGAACGTGTTCTCCGGAAGTCCCTTTGTGCTCTCCCGCTGGGATCCTGTGGATTT	23
Db	334	CTGTCTCTGCTTTTCCGTGGCACTCCCTTTGGGAAACCCATGATGGAACTCCCTGCTTT	392
QY	235	ACGAACCCGAGCCTGATCGCCCTGAGATTAATTGCGAAGACCACTCCCTACCCCTAT	297
Db	394	CTGCACCCGGAAGCCCAAGCAGCCCTTGTGTAAGGGGTGTGTAGACGTACAACTTCCAGGT	452
QY	295	TGTGCTCCAGAACTCAAGTGGCTGCTTTAGATCAACACATGCTCAAGGTGATTAAC	354
Db	454	GTCTGTACCAATATGTGGACACCTATACCAGGTTTAAAGGACCAAGATGTGGAAAC	512
QY	355	CCGAATTTGGAAGGTCCAAAGAATGCGCTCTACCTGAACATCTATAGCGCTGCGCAGCC	412
Db	514	CCCAACCGTAGCTGAGCAAGACGTGCTCTCACTCAACGTGTGGACAACATTAACCCCGG	572
QY	415	GATACAGAGTCCAAAGTCCCGCTTGTGGTGTGTTCCAGAGAGGCTCTTAAGACTGAC	474
Db	574	CTTACATCCCCCA---CCCTGTCTCTGTGTGATCTATAGGGGGTGGCTTTCAAGTGTGG	632
QY	475	TCAAGCTC-----CATCTTTGATGGGTCCGCTTGCTGCTATAGAGAGTGTGTT	528
Db	631	GCTCTCTCTTTGGACGTGACATGATGGCGCTTCTTGTGTACAGCGCAGAGAGACTGTGCTG	690
QY	529	GTTGTGTGTCAAGTACCGGCTAGAGAAATTTGGTTCTTCCAC---AATGGGATCACAT	585
Db	691	GTGTTCATATACTACCGGGTGGAGCTTTGGCTTCTGCGCTGCGGGAGCTCGAGAG	750
QY	586	GCTCCGGGAACTGGGCTTTCAAGACCAAGTGGCTGTCTGTCTCTGGTCCAGAAAGAC	644
Db	751	GGCCCGGGCAATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGTGCACTGGGTTCAGAAATAC	810
QY	646	ATCGAATTTTGTGGTGGGACCCCACTCTGTGACCAATCTTTGGCGAGTCCGCGGAGCC	705
Db	811	GTGGCAGCCTTGTGGGGGTGACCCGACATCAGTGAACGTGTTTGGGGAGAGCGCGGAGCC	870
QY	706	ATTAAGTTTTATAGCTTATACGTCTCCCAATGGCCAAAGCTTATTTCCAAAGCATC	765
Db	871	GGCTCGGTGGCAATGACTGCTGTGCCCGGACCGGGGCTGTTCACAGGGCCGTG	930

```

QY      766 ATGAGAGTGGGGTGCCA 784
          |||||
Db      931 CTCAGAGCGGTGCCCA 949

```

RESULT 15
US-08-370-156-1

Sequence 1, Application US/08370156
Patent No. 5932780
GENERAL INFORMATION:
APPLICANT: Sorreg, Hermona
APPLICANT: Zakut, Haim
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
ANTICHOLINESTERASE SUBSTANCES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reing, Ehnington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
Zip: 48099
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,156

TOPLOGY
US-08-370-156-1

Query Match	8.0%	Score 139.4;	DB 2;	Length 2256;
Best Local Similarity	55.1%	Pred. No. 5.8e-33;		
Matches 341; Conservative	0;	Mismatches 266;	Indels 12;	Gaps 3

QY 175 CCTGTGAACGTTTCTCTGGAGTCCCTTTCTGCTGCCCGGTGGAGATCCTCGATTT 23
Db 334 CCTGTCTCTGTCTTCTCTGGGCATCCCTTTTCGGAGCACCACCATGGAGCCCGCTGCTTT 393
QY 235 ACGAACCCGCAAGCTCGATCGCCCTGGGATTAACTTGGAGAGGACCTCTTACCCTAAT 294
Db 394 CTGCCACCGGAGCCCAAGACAGCTTTGTCAAGGGGTGTAGCGTTACACTTTCAGAGT 455
QY 295 TTGTGCTTCCGAACCTCAGGTGGCTGCTTTAGATCAACAATGCTCAAGGTGCATTAC 354
Db 454 GTCTCTACCAATATGTGACACCCCTAACCCAGATTTTGGAGGCACCGAGATGTGAAAC 513
QY 355 CCGAAATTCGGAGTGTCAAGAAATGCTGCTCTAAGTGAACATCTATGGCGCTGCCACGCC 414
Db 514 CCGAACCGGTAGCTGAGGGAAGAACTGCTGTACCTCAACGTGTGAAACACATATCCCCGG 573
QY 415 GATACAGGCTCAAGACTCCCGCTCTTGGTGTGTTCCAGAGGTGCTTCAAGACTGCG 474
Db 574 CCTTACATCCCCA----CCCTGTCTCTGTCTGAAATCTATGGGGGTGCTTCTACAGTGGG 630
QY 475 TCAGGCTC-----CATCTTGAATGGATCGGCGCTGGCTGCTATATGAGACGTGTGGTT 528

```
Db      631 GCCTCCTCTTGACGTGACGATGCGCCTTCTTGATACAGCCGAGAGACTGTGCTG 690
Oy      529 GTGTGCTCCAGTACCGGCTAGGAATTTGGTTCTTCACC--ACATGGATCAGCAT 585
Db      691 GTGTCAATGAACTACCGGGGTGGAGCCTTGGCTTCTGTGCCCTGCGGGGAGCCGAGAG 750
Oy      586 GCTCCGGGGAACCTGGGCTTCAAGGACAGGTGGCTGTCTGTCTGCTGGGTCCAGAGAAG 645
Db      751 GCCCGGGCAATGTGGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAAAC 810
Oy      646 ATCGAGTTCTTGGTGGGAGCCCAAGCTCTGTGACCAATCTTTGCGAGTCCGCGGAGCC 705
Db      811 GTGCGAGCCTTGGGGGTGACCCGACATCATGACGCTGTTGGGGAGAGCGGGAGGCC 870
Oy      706 ATAGTGTCTTCTAGTCTTACTGTCTCCCATGSCCAAGGCTTATTCACAAGCCATC 765
Db      871 GCCTGGGTGGGCATGCACTGTCTGTCCCGCCAGCCGGGGGCTGTTCACAGGGCCGTG 930
Oy      766 ATGAGAGTGGGGTGGCCA 784
Db      931 CTGCAGAGCGGTGCCCCCA 949
```

Search completed: June 14, 2005, 23:03:36
Job time : 296.702 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 937.402 Seconds
(without alignments)
11026.083 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 1746
Sequence: 1 atgcacagggagctacttc.....tttcttttgcctcttga 1746

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1746	100.0	2158	6	ABN84302
2	1746	100.0	2158	13	ABO89093
3	1652.8	94.7	1728	6	ABO86169
4	1651.6	94.6	2232	6	ABD40574
5	1639.6	93.9	2229	13	ACN42766
6	1631.2	93.4	2092	10	ADB62095
7	1435.6	82.2	1746	8	ABX72267
8	1433.4	82.0	1746	8	ABO86170
9	1386.2	77.4	1857	7	ADP19681
10	1344.4	77.0	1962	10	ADA53168
11	1099	62.9	2145	12	ADP50146
12	1063.2	60.9	1629	12	ADP50144
13	1021.4	58.5	1071	6	ABO86171
14	905.2	51.8	1244	10	ADC55523
15	634.6	36.3	921	6	ADP33344
16	508.6	29.1	965	6	ABZ11273
17	508.6	29.1	965	12	ADM43791
18	406.6	23.3	1680	8	AAZ25258
19	406.6	23.3	1680	8	AAZ52013
20	406.6	23.3	1699	8	ABZ24052

21	406.6	23.3	2169	8	ABZ24051	Abz24051	CES2 rela
22	406.6	23.3	2191	6	ABL62491	AbL62491	Colon ade
23	406.6	23.3	2191	6	ABL62492	AbL62492	Colon ade
24	406.6	23.3	2191	6	ABL63768	AbL63768	Breast ca
25	406.6	23.3	2191	6	ABN97359	AbN97359	Gene #385
26	406.6	23.3	2484	3	AAZ21884	AAZ21884	Human bre
27	405	23.2	1748	8	ABZ24055	AbZ24055	CES2 rela
28	400.2	22.9	2101	3	AAZ00680	AAZ00680	Human Hyd
29	400.2	22.9	3776	4	AAZ08405	AAZ08405	Human sec
30	400.2	22.9	3824	3	AAZ66341	AAZ66341	CDNA enco
31	400.2	22.9	3824	4	AAZ21514	AAZ21514	Human CDN
32	400.2	22.9	3824	6	ABK33640	ABK33640	CDNA enco
33	400.2	22.9	3824	6	ABL88204	ABL88204	Human PRO
34	400.2	22.9	3824	6	ABL95653	ABL95653	Human ang
35	400.2	22.9	3824	8	ACD28810	ACD28810	Human sec
36	400.2	22.9	3824	8	ACA03873	ACA03873	CDNA enco
37	400.2	22.9	3824	8	ABX89411	ABX89411	DNA enco
38	400.2	22.9	3824	8	ACA06084	ACA06084	CDNA enco
39	400.2	22.9	3824	8	ACD42065	ACD42065	Human sec
40	400.2	22.9	3824	8	ACA68601	ACA68601	Novel hum
41	400.2	22.9	3824	8	ACA04294	ACA04294	Human CDN
42	400.2	22.9	3824	9	ACA67707	ACA67707	CDNA enco
43	400.2	22.9	3824	9	ADA66060	ADA66060	Novel hum
44	400.2	22.9	3824	9	ADA76547	ADA76547	Novel hum
45	400.2	22.9	3824	9	ADA76491	ADA76491	Human PRO

ALIGNMENTS

RESULT 1	ABN84302	ABN84302	standard; CDNA; 2158 BP.
ID	ABN84302;		
XX			
AC	ABN84302;		
XX			
DT	23-SEP-2002	(first entry)	
XX			
DE	Human carboxylesterase family member 53010 coding sequence.		
XX			
KW	Carboxylesterase; enzyme; human; analgesic; nootropic; antiinflammatory;		
KW	diagnosis; gene therapy; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PH			
FT	Key	Location/Qualifiers	
FT	CDS	96..1841	
FT		/*tag= a	
FT		/product= "53010"	
FT		/note= "the CDS is also claimed in Claim 1"	
FT	sig_peptide	96..173	
FT		/*tag= b	
FT	mat_peptide	174..1838	
FT		/*tag= c	
PN	WO200250256-A2.		
XX			
PD	27-JUN-2002.		
XX			
PF	18-DEC-2001; 2001WO-US049075.		
XX			
PR	18-DEC-2000; 2000US-0256369P.		
XX			
PA	28-MAR-2001; 2001US-0279508P.		
XX			
PI	(MILL-) MILLENNIUM PHARM INC.		
XX			
DR	Curtis RAJ, Silos-Santiago I;		
XX			
DR	WPI, 2002-547936/58.		
XX			
PT	F-PSDB; ABB79537.		
XX			
PT	53010 nucleic acids, useful for diagnosing and treating e.g. vascular diseases, autoimmune diseases, or neurodegenerative diseases, as		

PT surrogate markers, in tissue typing and chromosome mapping.
XX
PS Claim 1; Page 109-110; 122pp; English.
XX

The present nucleotide sequence, the coding region of which is also claimed, is that of cDNA encoding human 53010, a novel member of the carboxylesterase family. The invention provides 53010 nucleic acids, antisense molecules, expression vectors, host cells, transgenic animals, 53010 proteins, fusion proteins, antigenic peptides, anti-53010 antibodies and methods for detecting the presence of 53010 polypeptides or nucleic acids, of identifying a compound that binds to the 53010 polypeptide, and of modulating the activity of the polypeptide. The 53010 nucleic acids and polypeptides can act as novel diagnostic and therapeutic agents for controlling disorders involving aberrant or deficient hydrolysis of carboxylic esters. As 53010 mRNA is highly expressed in the central and peripheral nervous system, and its expression is regulated in some rodent pain models, 53010 molecules can also act as novel diagnostic targets and therapeutic agents for controlling neurological disorders, such as pain-related disorders. A claimed method of treating or preventing a disorder (especially a pain-related disorder) characterised by aberrant activity of a 53010-expressing cell involves administering a compound that modulates 53010 activity or expression. 53010 nucleic acids are also useful in chromosome mapping, tissue typing, in forensic biology, prognostic assays, in arrays, for detection of variations or mutations, as surrogate markers and in pharmacogenomics

Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;

Query Match 100.0%; Score 1746; DB 6; Length 2158;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCACAGGAGCTTACTCATCTGCTTCAATGAGTCTTTTCTTGAATTCCTCAGCCC 60
DB 96 ATGCCACAGGAGCTTACTCATCTGCTTCAATGAGTCTTTTCTTGAATTCCTCAGCCC 155
QY 61 CTGTTGGGACACAGACAGTGGGAAAAAATGGGCCCTTCTGTGAAGGGCCACAGAGAAC 120
DB 156 CTGTTGGGACACAGACAGTGGGAAAAAATGGGCCCTTCTGTGAAGGGCCACAGAGAAC 215
QY 121 ACCAGGCTGGAGTGGATTCAGGGCAAGCAATCTGTGCTGGAGAGCCCTGTGCTGTG 180
DB 216 ACCAGGCTGGAGTGGATTCAGGGCAAGCAATCTGTGCTGGAGAGCCCTGTGCTGTG 275
QY 181 AACGTGTTCTCGAGTCCCTTTGTGCTCTCCCGCTGGAGTCCCTGGATTTTACGAAC 240
DB 276 AACGTGTTCTCGAGTCCCTTTGTGCTCTCCCGCTGGAGTCCCTGGATTTTACGAAC 335
QY 241 CCGGAGCTGCATCGCCCTGGAGTAACTTGGAGAAAGCACTCTCAATTTTGTGC 300
DB 336 CCGGAGCTGCATCGCCCTGGAGTAACTTGGAGAAAGCACTCTCAATTTTGTGC 395
QY 301 CTCGAGACTCAGAGTGGCTGCTTAGTCAACAATGCTCAAGAGTGCATTTACCCGAA 360
DB 396 CTCGAGACTCAGAGTGGCTGCTTAGTCAACAATGCTCAAGAGTGCATTTACCCGAA 455
QY 361 TTCCGAGTGCAGAAAGTGCCTTAAGCAATCTATGCGCTGCGCCACGCGGATACA 420
DB 456 TTCCGAGTGCAGAAAGTGCCTTAAGCAATCTATGCGCTGCGCCACGCGGATACA 515
QY 421 GGCTCCAAAGCTCCCTCTTGTGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAGCC 480
DB 516 GGCTCCAAAGCTCCCTCTTGTGTGTGTTCCAGAGAGGTGCTTCAAGACTGCTCAGCC 575
QY 481 TCCATCTTTGATGGGTCCGCTGAGTGCCTATGAGGACGCTGTTGTGTGTGTGTGTGT 540
DB 576 TCCATCTTTGATGGGTCCGCTGAGTGCCTATGAGGACGCTGTTGTGTGTGTGTGTGT 635
QY 541 TACCGGCTAGGAATATTTGTTCTTCAACCATGAGATCAGCATGCTCCGGGAACTGG 600
DB 636 TACCGGCTAGGAATATTTGTTCTTCAACCATGAGATCAGCATGCTCCGGGAACTGG 695

QY 601 GCCTTCAAGACCAAGTGGCTGCTCTGTCTGGGTCCAGAAACATCGAGTTCTTCGGT 660
DB 696 GCCTTCAAGACCAAGTGGCTGCTCTGTCTGGGTCCAGAAACATCGAGTTCTTCGGT 755
QY 661 GGGGAGCCAGGCTCTGTGACCAATCTTTGGCAGTCCCGGGAGCCATAGTCTTCTAGT 720
DB 756 GGGGAGCCAGGCTCTGTGACCAATCTTTGGCAGTCCCGGGAGCCATAGTCTTCTAGT 815
QY 721 CTTATAGTGTCTCCATGAGCCCAAGAGCTTATTCACAAAGCATATGAGAGTGGAGTG 780
DB 816 CTTATAGTGTCTCCATGAGCCCAAGAGCTTATTCACAAAGCATATGAGAGTGGAGTG 875
QY 781 GCCATCATCCCTTACCTGAGGAGCCATATTTAGAAAGTGAAGAGCTTCAGAGTGGT 840
DB 876 GCCATCATCCCTTACCTGAGGAGCCATATTTAGAAAGTGAAGAGCTTCAGAGTGGT 935
QY 841 GCACATTTCTGTGTTGATCAATAGCCGACAGCTCTGAGGAGCCGCTGAGGAGTGTAGACA 900
DB 936 GCACATTTCTGTGTTGATCAATAGCCGACAGCTCTGAGGAGCCGCTGAGGAGTGTAGACA 995
QY 901 AAACCTTCAAGAGAGTGTGACCTCAGCCAGAAAACAAAGTCTTTCATCTGAGTGGT 960
DB 996 AAACCTTCAAGAGAGTGTGACCTCAGCCAGAAAACAAAGTCTTTCATCTGAGTGGT 1055
QY 961 GATGGTCTTCTTCTTCAATAGAGCCCTTATATTTGTCTCAAGAAAGCTTTAAAGCA 1020
DB 1056 GATGGTCTTCTTCTTCAATAGAGCCCTTATATTTGTCTCAAGAAAGCTTTAAAGCA 1115
QY 1021 ATTCTTCATCATCGAGAGTCAATTAACACAGAGTGTGCTTCCGCTGCTCATGAAGAG 1080
DB 1116 ATTCTTCATCATCGAGAGTCAATTAACACAGAGTGTGCTTCCGCTGCTCATGAAGAG 1175
QY 1081 GCTCTGAGATCTCTCATGAGTCTCAACAAAGTCCCTTCCATCTGATTAACAAACATC 1140
DB 1176 GCTCTGAGATCTCTCATGAGTCTCAACAAAGTCCCTTCCATCTGATTAACAAACATC 1235
QY 1141 CTGACATCCCGCTCAGTATTTGGACCTTGTGGCTTATTAATCTTCATGACCAAGCAC 1200
DB 1236 CTGACATCCCGCTCAGTATTTGGACCTTGTGGCTTATTAATCTTCATGACCAAGCAC 1295
QY 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGGAGATGTCTTCTTGTGCTG 1260
DB 1296 TCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGGAGATGTCTTCTTGTGCTG 1355
QY 1261 CTTGACCTGATCAAGTGTGATATCAAGAGATGCTGTGACCTGTCTATCTTATGAG 1320
DB 1356 CTTGACCTGATCAAGTGTGATATCAAGAGATGCTGTGACCTGTCTATCTTATGAG 1415
QY 1321 TTTCGGACCGGCTCAGTGTCTTGAAGACAGAAAGCCGCTTTGTCAAAAGCCGACAC 1380
DB 1416 TTTCGGACCGGCTCAGTGTCTTGAAGACAGAAAGCCGCTTTGTCAAAAGCCGACAC 1475
QY 1381 GCTGATGAATCCGCTTGTGTGTGTGCTTCTCTGAAGGGGAGCATTTGTATGTTC 1440
DB 1476 GCTGATGAATCCGCTTGTGTGTGTGCTTCTCTGAAGGGGAGCATTTGTATGTTC 1535
QY 1441 GAAGAGCCACGAGAGGAGAGTAACTAGAGCCGGAAGATGATGAATAGTGGCTTAC 1500
DB 1536 GAAGAGCCACGAGAGGAGAGTAACTAGAGCCGGAAGATGATGAATAGTGGCTTAC 1595
QY 1501 TTGTCTGAAACCGGAATCTTAATGGAAAGACCTGTCTGTGAGCACTTAATATCTG 1560
DB 1596 TTGTCTGAAACCGGAATCTTAATGGAAAGACCTGTCTGTGAGCACTTAATATCTG 1655
QY 1561 ACTGACAGTAACTCTCAAGTGTGACTTGAACATGAGCTTGGACAGAGACTCAAGAACCG 1620
DB 1656 ACTGACAGTAACTCTCAAGTGTGACTTGAACATGAGCTTGGACAGAGACTCAAGAACCG 1715
QY 1621 CGGGTGAATTTTGTGACACAGCAATCCCTGATCTGTGCGCTCCGACATAGCTCCAC 1680
DB 1716 CGGGTGAATTTTGTGACACAGCAATCCCTGATCTGTGCGCTCCGACATAGCTCCAC 1775
QY 1681 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

Db 1056 GATGAGCTCTTCTTCCATAGAGCCTCTGATATATGTCTCAGAAAGATTAAAGCA 1115
Qy 1021 ATTCCCTCCATCAGAGAGTAAATACCAAGAGTGGCTTCTGCTGCTTAAAGAG 1080
Db 1116 ATTCCTTCATCAGTGGAGTCAATACCAAGAGTGGCTTCTGCTGCTTAAAGAG 1175
Qy 1081 GCTCCTAGATCCCTAGTGGCTCAAGAGTCCCTTCCCTCCATCGATACAAACATC 1140
Db 1176 GCTCCTAGATCCCTAGTGGCTCAAGAGTCCCTTCCCTCCATCGATACAAACATC 1235
Qy 1141 CTGCACTCCCGCTCAGTATTTGACCTTGTGCTAATGATACTTCACAGACAC 1200
Db 1236 CTGCACTCCCGCTCAGTATTTGACCTTGTGCTAATGATACTTCACAGACAC 1295
Qy 1201 TCCCTGCTGAAATCCGAGCAGCTCTGACCTTGTGCTAATGATACTTCACAGAC 1260
Db 1296 TCCCTGCTGAAATCCGAGCAGCTCTTGTGCTAATGATACTTCCTTGTGCTC 1355
Qy 1261 CCGTCACTGATCAAGCTCGATATCAAGAGATGTGTCACCTGTCTACTTCTATGAG 1320
Db 1356 CCGTCACTGATCAAGCTCGATATCAAGAGATGTGTCACCTGTCTACTTCTATGAG 1415
Qy 1321 TTTGGGACCGGCTCAGTCTTTGAAAGACAGAGCCGCTTTGTCAAAGCCGACAC 1380
Db 1416 TTTGGGACCGGCTCAGTCTTTGAAAGACAGAGCCGCTTTGTCAAAGCCGACAC 1475
Qy 1381 GCTGATGAAGTCCGCTTTGTGCTCGGAGGCTTCTGAAAGGGGACATTTGATGTC 1440
Db 1476 GCTGATGAAGTCCGCTTTGTGCTCGGAGGCTTCTGAAAGGGGACATTTGATGTC 1535
Qy 1441 GAAGAGCCACGAGAGAGAGAAATTACTAGCCGAGAGATGATAATCTGGCTACC 1500
Db 1536 GAAGAGCCACGAGAGAGAGAAATTACTAGCCGAGAGATGATAATCTGGCTACC 1595
Qy 1501 TTTGCTCGAAGCGGGAATCTATGAGGAGACGCTGTCTGTGGCCAGCTTATATCTG 1560
Db 1596 TTTGCTCGAAGCGGGAATCTATGAGGAGACGCTGTCTGTGGCCAGCTTATATCTG 1655
Qy 1561 ACTGAGCAGTACCTCAGCTGATCTTGAACATGAGCTCGACAGAGACTCAAGAACCG 1620
Db 1656 ACTGAGCAGTACCTCAGCTGATCTTGAACATGAGCTCGACAGAGACTCAAGAACCG 1715
Qy 1621 CGGCTGATTTTGGACACGACATCCCTGATCTGTCTGCTCCGACATGCTCAC 1680
Db 1716 CGGCTGATTTTGGACACGACATCCCTGATCTGTCTGCTCCGACATGCTCAC 1775
Qy 1681 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1776 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
Qy 1741 CCTTGA 1746
Db 1836 CCTTGA 1841

RESULT 3
ABQ86169
ID ABQ86169 standard; DNA; 1728 BP.
AC ABQ86169;
XX
XX 10-SEP-2002 (first entry)
XX
DE Novel human gene. SEQ ID 40.
XX
XX

Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW neurotrophic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antitumor; antitumor;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;

KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo-transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200250105-A1.
XX
XX 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
PA (SMK) SMITHKLINE BEECHAM CORP.
PA (SMK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Serum UC, Xie Q;
XX
XX WPI: 2002-508784/54.
DR P-PSDB; ABP61004.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 2(a); Page 251; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotrophic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antitumor, antitumor, antitumor, antitumor, antitumor,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic disease, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABQ86130-ABQ86184 represent novel human CDNA's
CC of the invention
XX
XX

SO Sequence 1728 BP; 367 A; 483 C; 437 G; 441 T; 0 U; 0 Other;

Query Match 94.7%; Score 1652.8; DB 6; Length 1728;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 91 GGAGCTTCTGTTGAAGGCGCACAGAGAACACAGCTGGAGTGTGATTCAGGCAAGCA 150


```
Db 73 GGGCTTCTGCTGGAAGGCGCAGAGAGAACACAGGCTGGAGATTCAGGGCAAGCAA 132
Qy 151 GTCAATGTCGTGGGAAGCCCTGCTGTGAACGTGTTCTCTGGAGTCCCTTTCTGT 210
Db 133 GTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTCTGT 192
Qy 211 CCCCCGTGGGATCCCTGCGATTTTACGAACCCGACGCTGATCCGCTGGATTAATTG 270
Db 193 CCCCCGTGGGATCCCTGCGATTTTACGAACCCGACGCTGATCCGCTGGATTAATTG 252
Qy 271 CGAAGAGCCACTCTCTACCTTAATTTGTGCTTCAAGACTCAGAGTGGCTGTCTTAAT 330
Db 253 CGAAGAGCCACTCTCTACCTTAATTTGTGCTTCAAGACTCAGAGTGGCTGTCTTAAT 312
Qy 331 CAACATGCTTCAGAGTGTATTAACCGGAATTCGGAGTGTAGGAAGATCTGCTCTACCT 390
Db 313 CAACATGCTTCAGAGTGTATTAACCGGAATTCGGAGTGTAGGAAGATCTGCTCTACCT 372
Qy 391 AACATCTATGCGCTGCGCCAGCGATACAGGCTCAAGGCTCCCGCTTGGTGTGTTTC 450
Db 373 AACATCTATGCGCTGCGCCAGCGATACAGGCTCAAGGCTCCCGCTTGGTGTGTTTC 432
Qy 451 CCAGAGGTCGCTTCAAGACTGCTCAAGCTTCATCTTTGATGGTTCGCTGCTGCTGC 510
Db 433 CCAGAGGTCGCTTCAAGACTGCTCAAGCTTCATCTTTGATGGTTCGCTGCTGCTGC 492
Qy 511 TATGAGACGTGCTGTGTTGTGTGCTCCAGTACCGGCTAAGAAATTTGTTTCTTCAAC 570
Db 493 TATGAGACGTGCTGTGTTGTGTGCTCCAGTACCGGCTAAGAAATTTGTTTCTTCAAC 552
Qy 571 ACATGGATTCAGCATGCTCCGGGAACTGGGCTTCAAGGACAGAGTGGCTGTCTGTGTC 630
Db 553 ACATGGATTCAGCATGCTCCGGGAACTGGGCTTCAAGGACAGAGTGGCTGTCTGTGTC 612
Qy 631 TGGGTCCAGAAAGACATGAGTTCTTCGGTGGGAAACCCAGCTGTGTGACATCTTTGGC 690
Db 613 TGGGTCCAGAAAGACATGAGTTCTTCGGTGGGAAACCCAGCTGTGTGACATCTTTGGC 672
Qy 691 GAGTCCGCGGAGGACATTAAGTGTCTTAAGTCTTAATCTGTCTCCATGCGCAAGGCTTA 750
Db 673 GAGTCCGCGGAGGACATTAAGTGTCTTAAGTCTTAATCTGTCTCCATGCGCAAGGCTTA 732
Qy 751 TTCCCAAGGACCATGAGAGAGTGGGAGTGGCCATCATCCCTTACCTGGAGGCCATGAT 810
Db 733 TTCCCAAGGACCATGAGAGAGTGGGAGTGGCCATCATCCCTTACCTGGAGGCCATGAT 792
Qy 811 TATGAGAGAGTGAAGACTGCGAGGTGTGACATTTCTGTGTAAACATGCTGCAGAC 870
Db 793 TATGAGAGAGTGAAGACTGCGAGGTGTGACATTTCTGTGTAAACATGCTGCAGAC 852
Qy 871 TCTGAGGCTCTGTGAGGTGCTGAGACAAAAACCTTCAGAGAGCTGTGACCTCAAC 930
Db 853 TCTGAGGCTCTGTGAGGTGCTGAGACAAAAACCTTCAGAGAGCTGTGACCTCAAC 912
Qy 931 CAGAAACAAAGTCTTTTCACTCGAGGTGTGATGAGTGTCTTTCTCTAATAGCCTCA 990
Db 913 CAGAAACAAAGTCTTTTCACTCGAGGTGTGATGAGTGTCTTTCTCTAATAGCCTCA 972
Qy 991 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTCCTTCATCATGAGAGTCAATACAC 1050
Db 973 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTCCTTCATCATGAGAGTCAATACAC 1032
Qy 1051 GAGTGTGCTTCTGTGCTGCTTAAGAGAGGCTCTGAGATCTCAAGTGTCCCAACAG 1110
Db 1033 GAGTGTGCTTCTGTGCTGCTTAAGAGAGGCTCTGAGATCTCAAGTGTCCCAACAG 1092
Qy 1111 TCCCTTGGCTTCATCTGATTAAGAAACATCCGAGCATCCCGGCTCAAGATTTGACCTT 1170
Db 1093 TCCCTTGGCTTCATCTGATTAAGAAACATCCGAGCATCCCGGCTCAAGATTTGACCTT 1152
Qy 1171 GTGGCTAAATGAAATCTTCATGACAAAGACTCCCTGACTGAATCCGAGACAGTCTTCTG 1230
```

```
Db 1153 GTGGCTAAATGAAATCTTCATGACAAAGACTCCCTGACTGAATCCGAGACAGTCTTCTG 1212
Qy 1231 GACTGTCTTGGAGATGTGTTCTTTGTGCTGCTGACATGATCAAGTCAATACAGA 1290
Db 1213 GACTGTCTTGGAGATGTGTTCTTTGTGCTGCTGACATGATCAAGTCAATACAGA 1272
Qy 1291 GATGTGTGTGACCTGTCTACTTCTATGAGTTTGGGACCCGCTCAAGTCTTTGAAG 1350
Db 1273 GATGTGTGTGACCTGTCTACTTCTATGAGTTTGGGACCCGCTCAAGTCTTTGAAG 1332
Qy 1351 ACGAAGCCGCTTTTGTGTAAAGCCGACCAAGTGTAAAGTCCGCTTGTGTGCTGCT 1410
Db 1333 ACGAAGCCGCTTTTGTGTAAAGCCGACCAAGTGTAAAGTCCGCTTGTGTGCTGCT 1392
Qy 1411 GCTTCTTGAAGGGGAGCATTTGTTATGTTTGAAGAGCCAGAGAGAGAGAGTTACTG 1470
Db 1393 GCTTCTTGAAGGGGAGCATTTGTTATGTTTGAAGAGCCAGAGAGAGAGAGTTACTG 1452
Qy 1471 AGCCGGAAGATGATGAATTAATCTGGGCTACCTTGTCTGACCCGGAATCTTAATGGAA 1530
Db 1453 AGCCGGAAGATGATGAATTAATCTGGGCTACCTTGTCTGACCCGGAATCTTAATGGAA 1512
Qy 1531 GACTGTCTGTGTGCTGCTTATATCTGACTAGACATCTCAAGTGTGAAC 1590
Db 1513 GACTGTCTGTGTGCTGCTTATATCTGACTAGACATCTCAAGTGTGAAC 1572
Qy 1591 ATGAGCTTCGAGACAGAGCTCAAGAAACCCGGGGGTGATTTTGGACAGACCATCC 1650
Db 1573 ATGAGCTTCGAGACAGAGCTCAAGAAACCCGGGGGTGATTTTGGACAGACCATCC 1632
Qy 1651 CTGATCTGTGTGCTTCCGACATGCTCCAGAGTCTCTTCTTCTTAATCTTCTCT 1710
Db 1633 CTGATCTGTGTGCTTCCGACATGCTCCAGAGTCTCTTCTTCTTAATCTTCTCT 1692
Qy 1711 CTCTCTCAGCCTTTCTTTTCTTTTGTGCTCTTGA 1746
Db 1693 CTCTCTCAGCCTTTCTTTTCTTTTGTGCTCTTGA 1728
```

RESULT 4

AAd40574 standard; cDNA, 2232 BP.

AAd40574;

30-OCT-2002 (first entry)

Human drug metabolizing enzyme (DME-10) cDNA.

Human; drug metabolizing enzyme; autoimmune; inflammatory disorder;
acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;
proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;
asthma; neurological disorder; Alzheimer's disease; Huntington's disease;
dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;
drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;
renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;
anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;
gout; gastrointestinal disorder; gene therapy; virulence; liver disorder;
anticonvulsant; nootropic; enzyme; DME-10; gene; ss.

Homo sapiens.

Location/Qualifiers

1..1929

/*tag= a

/*product= "Human DME-10"

sig_peptide 1..150

mat_peptide 151..1926

/*tag= c

/*product= "Mature human DME-10"

WO200246426-A2.

XX 13-JUN-2002.
PD
XX
PF 04-DEC-2001; 2001MO-US047429.
XX
PR 08-DEC-2000; 2000US-0254308P.
XX
PR 15-DEC-2000; 2000US-0256189P.
XX
PR 21-DEC-2000; 2000US-0257713P.
PR 19-JAN-2001; 2001US-0262706P.
PR 02-FEB-2001; 2001US-0266020P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Sanjwalala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,
PI Lee BA, Ding L, Hafalia AJD, Tang YT, Yue H, Tribouley CM, Lu DAM,
PI Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AK, Lu Y,
PI Ison CH;
XX
XX WPI; 2002-519668/55.
DR P-PSDB; AAB25025.
XX
XX Novel human drug metabolizing polypeptide, useful in diagnosis,
PT prevention or treatment of autoimmune/inflammatory, cell proliferative,
PT neurological, developmental, endocrine, metabolic and gastrointestinal
PT disorders.
XX
XX Claim 78; Page 167-168; 169pp; English.
XX
XX The invention relates to an isolated human drug metabolizing enzyme (DME)
XX and its nucleotide. DME is useful for diagnosing, treating or preventing
XX disorders associated with aberrant expression of DME, where the disorders
XX are selected from autoimmune/inflammatory disorder such as acquired
XX immunodeficiency syndrome (AIDS), asthma, atherosclerosis, psoriasis,
XX uveitis; a cell proliferative disorder such as arteriosclerosis,
XX cirrhosis, hepatitis, and cancer; a neurological disorder such as
XX Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;
XX a developmental disorder such as renal tubular acidosis, epilepsy,
XX anaemia; an endocrine disorder such as adenoma, thrombosis and infections
XX in an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic
XX disorder such as cystic fibrosis, diabetes and goitre; a gastrointestinal
XX disorder such as anorexia, peptic ulcer, and liver disorders. DME is
XX useful in a number of drug screening techniques and to analyse the
XX proteome of a tissue or cell type. The invention is useful for creating
XX knock-in humanised animals or transgenic animals to model human diseases,
XX in somatic or germline gene therapy, to generate a transcript image of a
XX tissue or cell type, for detecting differences in the chromosomal
XX location due to translocation, inversion, etc. among normal, carrier or
XX affected individuals, and as hybridisation probes for mapping naturally
XX occurring genomic sequences. The present sequence is human DME-10 cDNA
XX
XX Sequence 2232 BP; 504 A; 586 C; 549 G; 593 T; 0 U; 0 Other;
XX
XX Query Match 94.6%; Score 1651.6; DB 6; Length 2232;
XX Best Local Similarity 99.5%; Pred. No. 0;
XX Matches 1657; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX
QY 81 GGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGAAACAGAGCTGGAGTGAATCA 140
DB 264 GGTCAAGAGAGGGCTTCTGCTGAAGGCCACAGAGAAACAGAGCTGGAGTGAATCA 323
QY 141 GGGCAAGAGTCACTGTGCTGGGAAGCCCTGTGAGAGCTGTTCCTCGAGTCCC 200
DB 324 GGGCAAGAGTCACTGTGCTGGGAAGCCCTGTGAGAGCTGTTCCTCGAGTCCC 383
QY 201 CTTTGCTGCTCCCGCTGGAGCTCTGCGATTTACGAACCCGAGCTGATCGCCCTG 260
DB 384 CTTTGCTGCTCCCGCTGGAGCTCTGCGATTTACGAACCCGAGCTGATCGCCCTG 443
QY 261 GGATTAATCTGGAGAACCCATCTTAACCTAATTTGTGCTTCAGAACTCAAGTGGCT 320
DB 444 GGATTAATCTGGAGAACCCATCTTAACCTAATTTGTGCTTCAGAACTCAAGTGGCT 503
QY 321 GCTCTAGATCAACATGCTCAAGGTGCAATTACCCGAATTGGAGTGTCAAGAACTG 380

DB 504 GCTCTTAGATCAACATGCTCAAGTGTGATTAACCGAAATTGGAGTGTCAAGAACTG 563
QY 381 CCTTAACCTGAACATCTAATGGCCCTGCCACGCCCATACAGGCTTCAGAGCTCCCGTCTT 440
DB 564 CTTTACCTGAACATCTAATGGCCCTGCCACGCCCATACAGGCTTCAGAGCTCCCGTCTT 623
QY 441 GGTGTGGTTCCAGAGAGGTGCTTCAGAACTGAGCTCAAGCTTCATCTTTGATGGGTCCG 500
DB 624 GGTGTGGTTCCAGAGAGGTGCTTCAGAACTGAGCTCAAGCTTCATCTTTGATGGGTCCG 683
QY 501 CCTGCTGCTCATAGAGACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
DB 684 CTTGCTGCTCATAGAGACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743
QY 561 TTTCTTCAACCAATAGGATTCAGATGCTCGGGGAACTGGGCTTCAGAAAGCCAGGTGGC 620
DB 744 TTTCTTCAACCAATAGGATTCAGATGCTCGGGGAACTGGGCTTCAGAAAGCCAGGTGGC 803
QY 621 TGCTCTGCTGGGTTCAGAAAGAAATCGAGATTCTGCTGGGGAGCCCAAGCTCTGTGAC 680
DB 804 TGCTCTGCTGGGTTCAGAAAGAAATCGAGATTCTGCTGGGGAGCCCAAGCTCTGTGAC 863
QY 681 CATCTTTGGGAGTCCCGGGAGCCATTAAGTTTCTAAGTTTCTAAGTTTCTAAGTTTCTAAGTT 740
DB 864 CATCTTTGGGAGTCCCGGGAGCCATTAAGTTTCTAAGTTTCTAAGTTTCTAAGTTTCTAAGTT 923
QY 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGGA 800
DB 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGGA 983
QY 801 GGCCCAATGATTATGAGAAAGTGAAGACCTGCAAGTGTTCACATTTCTGTGGTAAACA 860
DB 984 GGCCCAATGATTATGAGAAAGTGAAGACCTGCAAGTGTTCACATTTCTGTGGTAAACA 1043
QY 861 TGCTCTGAGCTTGAAGCCCTGCTGAGAGTGCCTGAGGACAAACCTCCCAAGAGAGTGTCT 920
DB 1044 TGCTCTGAGCTTGAAGCCCTGCTGAGAGTGCCTGAGGACAAACCTCCCAAGAGAGTGTCT 1103
QY 921 GACCTCAGCCAGAAAACAAAGTCTTTCATCTGAGTGTGTGATGTGTCTTCTTCTTA 980
DB 1104 GACCTCAGCCAGAAAACAAAGTCTTTCATCTGAGTGTGTGATGTGTCTTCTTCTTA 1163
QY 981 TGAGCCCTTAAGTCTAATGTGTCTAGAAAGATTTAAAGCAATTCCTTCATCTCGAAGT 1040
DB 1164 TGAGCCCTTAAGTCTAATGTGTCTAGAAAGATTTAAAGCAATTCCTTCATCTCGAAGT 1223
QY 1041 CAATTAACAGAGTGTGTCTTCTGCTGCTATGAAGAGAGCTCTGAGATCTCTAGTGG 1100
DB 1224 CAATTAACAGAGTGTGTCTTCTGCTGCTATGAAGAGAGCTCTGAGATCTCTAGTGG 1283
QY 1101 CTCGCAACAAGTCCCTGCTCCATCTGATTAACAAAACATCTTGACATCCGCTCAGTA 1160
DB 1284 CTCGCAACAAGTCCCTGCTCCATCTGATTAACAAAACATCTTGACATCCGCTCAGTA 1343
QY 1161 TTTGCACTTTGGGCTTAATGAATTAATCTTCATGACAAAGCACTCCGATCGAAATCCGAGA 1220
DB 1344 TTTGCACTTTGGGCTTAATGAATTAATCTTCATGACAAAGCACTCCGATCGAAATCCGAGA 1403
QY 1221 CAGTCTTGTGAACCTTGCTTGAAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280
DB 1404 CAGTCTTGTGAACCTTGCTTGAAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1463
QY 1281 ATATACAGAGATGCTGTGACCTGTCTAATCTTAATGATTTGGGACCCGCTCAAGTG 1340
DB 1464 ATATACAGAGATGCTGTGACCTGTCTAATCTTAATGATTTGGGACCCGCTCAAGTG 1523
QY 1341 CTTTGAAGACAGAAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTGT 1400
DB 1524 CTTTGAAGACAGAAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTGT 1583
QY 1401 GTTCGGTGTGCTTCTCTGAAGGGGACATTTGTTATGTTGAAAGAGCCAGAGAGAGA 1460


```

QY 921 GACCTGAGCCAGAAAAGAGTCTTCACTGAGTGGTGGATGAGTCTTCTTCTAA 980
DB 1103 GACCTGAGCCAGAAAAGAGTCTTCACTGAGTGGTGGATGAGTCTTCTTCTAA 1162
QY 981 TGAGCCTCTGAGTCTTATGCTCTGAGAAAGCAATTTAAAGCAATCTTCACTGAGT 1040
DB 1163 TGAGCCTCTGAGTCTTATGCTCTGAGAAAGCAATTTAAAGCAATCTTCACTGAGT 1222
QY 1041 CAATTAACAGAGTGGTCTTCCGCTGAGTAAAGAGGCTCCGAGATCTTCACTGAGT 1100
DB 1223 CAATTAACAGAGTGGTCTTCCGCTGAGTAAAGAGGCTCCGAGATCTTCACTGAGT 1282
QY 1101 CTCCAACAGTCCCTTGCCTTCATCTGATACAAACATCTGCACTATCCGCTCAGTA 1160
DB 1283 CTCCAACAGTCCCTTGCCTTCATCTGATACAAACATCTGCACTATCCGCTCAGTA 1342
QY 1161 TTTCGACCTTGTGGCTTAATTAATCTTCCATGACAAAGCACTCCCTGACTGAAATCCGAGA 1220
DB 1343 TTTCGACCTTGTGGCTTAATTAATCTTCCATGACAAAGCACTCCCTGACTGAAATCCGAGA 1402
QY 1221 CAGTCTTCTGAGTCTGCTGAGAGATGTTCTTGTGTCCGCTGACTGATCACTGAGT 1280
DB 1403 CAGTCTTCTGAGTCTGCTGAGAGATGTTCTTGTGTCCGCTGACTGATCACTGAGT 1462
QY 1281 AATTCACAGAGATGCTGAGTCACTTCTTATGAGTTCGACACCGGCTCAGTG 1340
DB 1463 AATTCACAGAGATGCTGAGTCACTTCTTATGAGTTCGACACCGGCTCAGTG 1522
QY 1341 CTTTGAAGACAGAGACCGGCTTGTCTCAAGCCGACCGCTGATGAATCCGCTTGT 1400
DB 1523 CTTTGAAGACAGAGACCGGCTTGTCTCAAGCCGACCGCTGATGAATCCGCTTGT 1582
QY 1401 GTTCGCTGTGCTTCTCTGAGAGGGGACATTTGATGTTCAAGAGAGCCAGAGAGAGA 1460
DB 1583 GTTCGCTGTGCTTCTCTGAGAGGGGACATTTGATGTTCAAGAGAGCCAGAGAGAGA 1642
QY 1461 GAACTTACAGAGCCGAGAAATGATGAATATCTGGGCTACCTTGTCTGAAACCGGAAATCC 1520
DB 1643 GAACTTACAGAGCCGAGAAATGATGAATATCTGGGCTACCTTGTCTGAAACCGGAAATCC 1702
QY 1521 TAATGGAGACGACTGTCTGTGGCAGCTTATATCTGACTGAGAGATCTTCACT 1580
DB 1703 TAATGGAGACGACTGTCTGTGGCAGCTTATATCTGACTGAGAGATCTTCACT 1762
QY 1581 GAGCTTGAACATGAGCTCTGAGACAGAGACTCAAGAACCGCGGTGATTTTGGACAG 1640
DB 1763 GAGCTTGAACATGAGCTCTGAGACAGAGACTCAAGAACCGCGGTGATTTTGGACAG 1822
QY 1641 CACCATCCCGCTGAGTCTGTGCTCCGACATGCTCCAGATCTTCTTCTTCTTAAC 1700
DB 1823 CACCATCCCGCTGAGTCTGTGCTCCGACATGCTCCAGATCTTCTTCTTCTTAAC 1882
QY 1701 TTTCCTCTCTCTCTGAGACCTTCTTCTTCTTGTGCTCTTGA 1746
DB 1883 TTTCCTCTCTCTCTGAGACCTTCTTCTTCTTGTGCTCTTGA 1928

```

RESULT 6
ADB62095
ID ADB62095 standard; cDNA; 2092 BP.

ADB62095;

04-DEC-2003 (first entry)

Human cDNA encoding clone BRAWH20021910.

Human; 98; gene; pharmaceutical; diagnostic; gene therapy;
tissue regeneration; cell regeneration; membrane protein;
signal transduction-related protein; transcription-related protein;
osteoporosis; neurological disease; cancer; tumour.

OS Homo sapiens.
XX Key Location/Qualifiers
XX CDS 394..1803
XX FT /tag= a
XX FT /product= "Clone BRAWH20021910 protein"
XX PN BP1308459-A2.
XX 07-MAY-2003.
XX 28-MAR-2002; 2002EP-00007401.
XX 05-NOV-2001; 2001JP-00379298.
XX 25-JAN-2002; 2002US-00350978.
XX (HELI-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y,
XX WPI; 2003-450961/43.
XX P-PSDB; ADB64065.
XX
XX New polynucleotides and polypeptides, useful for developing a diagnostic
XX marker or medicines for regulation of their expression and activity, or
XX as targets of gene therapy.
XX
XX Claim 1; Page; 222pp; English.

The invention discloses a polynucleotide comprising a sequence selected
from 1970 fully defined nucleotide sequences which encode novel
polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
or its partial peptide, an antibody binding to the polypeptide or peptide
of the polynucleotide, immunologically assaying the polypeptide or
peptide of the polynucleotide by contacting the polypeptide or peptide
with the antibody of the encoded protein, and observing the binding
between the two, a transformant carrying the polynucleotide in an
expressible manner and an antisense polynucleotide. The oligonucleotide
is useful as a primer for synthesizing the polynucleotide, or as a probe
for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX

Sequence 2092 BP; 464 A; 552 C; 497 G; 579 T; 0 U; 0 Other;

Query Match 93.4%; Score 1631.2; DB 10; Length 2092;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1653; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

```

QY 91 GGGCTTCTGCTGAGAGGACACAGAGAGACAGAGCTGGATGATGAGGAGCA 150
DB 138 GGGCTTCTGCTGAGAGGACACAGAGAGACAGAGCTGGATGATGAGGAGCA 197
QY 151 GTACTGTGCTGGAGAGCCCTGTGCTGAGAGTGTCTCTGAGATCCCTTTGCTGT 210
DB 198 GTCACTGTGCTGGAGAGCCCTGTGCTGAGAGTGTCTCTGAGATCCCTTTGCTGT 257
QY 211 CCCCCGCTGGAGATCCCTGAGATTAAGAACCCGAGCCCTGATCCGCTGGAGATACTTG 270
DB 258 CCCCCGCTGGAGATCCCTGAGATTAAGAACCCGAGCCCTGATCCGCTGGAGATACTTG 317

```

271 CGAAGAGCACTCTCCCTAATTT-----GTGCTCAAGAACTCAGAGTGT 320
318 CGAAGAGCACTCTCCCTAATTTTGAAGACAGAGTGTCTCCAGAACTCAGAGTGT 377
321 GCTCTTAATCAACAATGCTCAAGGTGATTAACCGAATTCGAGTGTCAAGAACTG 380
378 GCTCTTAATCAACAATGCTCAAGGTGATTAACCGAATTCGAGTGTCAAGAACTG 437
381 CCTTAACCTGAACATCTATGCGCTGCGCCAGCCGATACAGGCTCCAGCTCCCGTCTT 440
428 CCTTAACCTGAACATCTATGCGCTGCGCCAGCCGATACAGGCTCCAGCTCCCGTCTT 497
441 GGTGTGTTCCTCAAGAGGTGCTTCAAGACATGAGCTCAGCTCATTCTTGAATGGTCCG 500
498 GGTGTGTTCCTCAAGAGGTGCTTCAAGACATGAGCTCAGCTCATTCTTGAATGGTCCG 557
501 CCTGCTGCTATGAGAGAGT 560
558 CTTGCTGCTATGAGAGAGT 617
561 TTTCTTCAACATGAGAGT 620
618 TTTCTTCAACATGAGAGT 677
621 TGTCTGTCTGT 680
678 TGTCTGTCTGT 737
681 CATCTTTGGAGAGT 740
738 CATCTTTGGAGAGT 797
741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 800
798 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 857
801 GGGCCATATTAATGAGAGAGT 860
858 GGGCCATATTAATGAGAGAGT 917
861 TGGCTCAGACTGTGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 920
918 TGGCTCAGACTGTGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 977
921 GACCTTCAGCCAGAAACAAAGCTTTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGT 980
978 GACCTTCAGCCAGAAACAAAGCTTTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1037
981 TGAAGCTCTAATCTATGTGTCTCAAGAAAGCTTTTAAAGCAATTCCTTCAATCCGAGT 1040
1038 TGAAGCTCTAATCTATGTGTCTCAAGAAAGCTTTTAAAGCAATTCCTTCAATCCGAGT 1097
1041 CAATAACCAAGAGT 1100
1098 CAATAACCAAGAGT 1157
1101 CTCCAACAAGTCTCTTGGCTCTCATCTGATACAAACATCTGACATCCGCTCAGTA 1160
1158 CTCCAACAAGTCTCTTGGCTCTCATCTGATACAAACATCTGACATCCGCTCAGTA 1217
1161 TTTGCACTGT 1220
1218 TTTGCACTGT 1277
1221 CAGTCTTCTGAGCTGT 1280
1278 CAGTCTTCTGAGCTGT 1337
1281 AATATCAAGAGT 1340
1338 AATATCAAGAGT 1397

QY 1341 CTTTGAAGACAGAAAGCCGCTTTTGTCAAAACCGACAGCTGATGAAGTCCGCTTGT 1400
DB 1398 CTTTGAAGACAGAAAGCCGCTTTTGTCAAAACCGACAGCTGATGAAGTCCGCTTGT 1457
QY 1401 GTTCGAGTGTGTCTTCTGAAAGGAGACATTTGTATGTTGAAAGACCAAGAGAGGA 1460
DB 1458 GTTCGAGTGTGTCTTCTGAAAGGAGACATTTGTATGTTGAAAGACCAAGAGAGGA 1517
QY 1461 GAAGTTACTGAGCCGAGAAATGATAAATACGAGGCTACCTTGTCTGAAACGGGGAATCC 1520
DB 1518 GAAGTTACTGAGCCGAGAAATGATAAATACGAGGCTACCTTGTCTGAAACGGGGAATCC 1577
QY 1521 TAATGGAAAGCACTGTCTGTGTGACAGCTTATATCTGACTGAGCACTACCTCAGCT 1580
DB 1578 TAATGGAAAGCACTGTCTGTGTGACAGCTTATATCTGACTGAGCACTACCTCAGCT 1637
QY 1581 GGAATTGAACATGAGCTGTGTGACAGACTCAAGAAACCGGGGTGAGATTTTGAACAG 1640
DB 1638 GGAATTGAACATGAGCTGTGTGACAGACTCAAGAAACCGGGGTGAGATTTTGAACAG 1697
QY 1641 CACCATCCCCCTGATCTGTGTGCTCCGACATGCTCCACAGTCTCTTCTCTTAC 1700
DB 1698 CACCATCCCCCTGATCTGTGTGCTCCGACATGCTCCACAGTCTCTTCTCTTAC 1757
QY 1701 TTTCTCTCTCTCTCTCAGACCTTTCTTTTCTTTGTGTCTCTGA 1746
DB 1758 TTTCTCTCTCTCTCTCAGACCTTTCTTTTCTTTGTGTCTCTGA 1803

RESULT 7
ABX72267
ID ABX72267 strand: cDNA; 1746 BP.

XX ABX72267;
DT 03-JUN-2003 (first entry)

XX Human NOVX polynucleotide #98.

XX Human; NOVX; gene; 89; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subclavian stenosis; ventricular septal defect; VSD;
KW tuberosclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.

OS Homo sapiens.

XX W0200281498-A2.

XX 17-OCT-2002.

PF 03-APR-2002; 2002MO-US010780.

XX 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.

PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324802P.
PR 25-SEP-2001; 2001US-0324802P.
PR 17-OCT-2001; 2001US-0325684P.
PR 14-NOV-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
(CURA-) CURAGEN CORP.

Guo X, Kekuda R, Miller CE, Malyankar UM, Szytek KA;
Paturajan M, Liu X, Gusev VY, Li U, Vernet CM, Zehrhusen BD,
Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V,
Padigan W, Shimkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W;
Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ,
MacDougall JR, Rothenberg ME, Mazur A, Miller I, Peyman JA;
Ellerman K;

DR MPI: 2003-046858/04.
DR P-PSDB; AB054639.

PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.

PS Claim 17; Page 303; 666pp; English.

XX The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX
CC polynucleotides of the invention

XX Sequence 1746 BP; 402 A; 457 C; 459 G; 428 T; 0 U; 0 Other;

Query Match 82.2%; Score 1435.6; DB 8; Length 1746;
Best Local Similarity 95.3%; Pred. No. 0;

Matches 1496; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

QY 81 GGGAAAAAAGTGGGCTTCTGCTGAAGGGCCACAGAGAAACAGAGGCTGGGATGATTTCA 140
DB 69 GGTCAAGAGAGGGGCTTCTGCTGAAGGGCCACAGAGAAACAGAGGCTGGGATGATTTCA 128
QY 141 GGGCAAGACAGTCACTGTGCTGGGAAGCCCTGTGTCGTAACGTTCTCTCGAGTCCC 200
DB 129 GGGCAAGACAGTCACTGTGCTGGGAAGCCCTGTGTCGTAACGTTCTCTCGAGTCCC 188
QY 201 CTTTGCTGCTCCCGGCTGGGATCCCTGCGCATTTTACGAACCCGAGCTGTGATCGCCCTG 260
DB 189 CTTTGCTGCTCCCGGCTGGGATCCCTGCGCATTTTACGAACCCGAGCTGTGATCGCCCTG 248
QY 261 GGATTAACCTTGAGAGAACCACTCTCAACCTTAATTGTGCTTCAGAACTCAGAGTGGCT 320

DB 249 GGATTAACCTTGAGAGAACCACTCTCAACCTTAATTGTGCTTCAGAACTCAGAGTGGCT 308
QY 321 GCTCTTAAGTCAACACATGCTCAAGTGCATTAACCCGAATTCGAGTGTGAGAGACTG 380
DB 309 GCTCTTAAGTCAACACATGCTCAAGTGCATTAACCCGAATTCGAGTGTGAGAGACTG 368
QY 381 CCTTAACCTGAACATCTATGCGCTTCCGACGCGATACAGGCTTCAAGCTCCCGCTT 440
DB 369 CCTTAACCTGAACATCTATGCGCTTCCGACGCGATACAGGCTTCAAGCTCCCGCTT 428
QY 441 GGTGTGTTCCAGAGAGGCTTCAAGAGTGGCTCAAGCTTCAAGCTTCAAGTGGCTCCG 500
DB 429 GGTGTGTTCCAGAGAGGCTTCAAGAGTGGCTCAAGCTTCAAGCTTCAAGTGGCTCCG 488
QY 501 CCTGCTCTATAGAGAGCGT 560
DB 489 CCTGCTCTATAGAGAGCGT 548
QY 561 TTTCTTCAACACATGAGATTCAGATGCTTCCGAGGAACTGGGCTTCAAGACAGGTGGC 620
DB 549 TTTCTTCAACACATGAGATTCAGATGCTTCCGAGGAACTGGGCTTCAAGACAGGTGGC 608
QY 621 TGCTCTGTCTGGGTCCAGAGAAACATGAGATTCTTCGCTGGGAGCCAGCTCTGTGAC 680
DB 609 TGCTCTGTCTGGGTCCAGAGAAACATGAGATTCTTCGCTGGGAGCCAGCTCTGTGAC 668
QY 661 CATCTTTGGCGAGTCCCGGGAGCCATAGTGTCTTCACTTCACTTCACTTCACTTCACT 740
DB 669 CATCTTTGGCGAGTCCCGGGAGCCATAGTGTCTTCACTTCACTTCACTTCACTTCACT 728
QY 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTCACTTCACT 800
DB 729 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTCACTTCACT 788
QY 801 GGCCCATGATTATGAGAGAGTGAAGACCTGACAGTGTGACATTTCTGTGTGAACAA 860
DB 789 GGCCCATGATTATGAGAGAGTGAAGACCTGACAGTGTGTGACATTTCTGTGTGAACAA 848
QY 861 TGCTCTGAGACTGAGGCGCTTCTGAGTGTGCTGAGGACAAAACCTTCAAGAGACTGCT 920
DB 849 TGCTCTGAGACTGAGGCGCTTCTGAGTGTGCTGAGGACAAAACCTTCAAGAGACTGCT 908
QY 921 GACCTCAGCCAGAAAACAAAGCTTTTCACTGAGTGTGATGAGTGTCTTCTTCTTA 980
DB 909 GACCTCAGCCAGAAAACAAAGCTTTTCACTGAGTGTGATGAGTGTCTTCTTCTTA 968
QY 981 TAGCCTCTAGTCTATTTGTCTCAAGAAAGATTAAAGCAATTCCTTCATCATCGAGT 1040
DB 969 TAGCCTCTAGTCTATTTGTCTCAAGAAAGATTAAAGCAATTCCTTCATCATCGAGT 1028
QY 1041 CAATTAACCAAGAGTGTGCTTCTGAGTGTGCTTCAAGAGAGGCTTCTGAGTGTGCT 1100
DB 1029 CAATTAACCAAGAGTGTGCTTCTGAGTGTGCTTCAAGAGAGGCTTCTGAGTGTGCT 1088
QY 1101 CTCACCAAGTCCCTGCTCCATCTGATACAA-----ACATCTCTGA 1145
DB 1089 CACTCCCTCAACCGAGTGAAGCACTTTGGCTTCAACAGCTGGGCACTTCCAGAGAGCA 1148
QY 1146 CATCCGCGCTCAGTATTTGCACTTTGTGCTTAATGAATCTTCATGACAGACATCCCT 1205
DB 1149 TATCCGCGCTCAGTATTTGCACTTTGTGCTTAATGAATCTTCATGACAGACATCCCT 1208
QY 1206 GACTGAATCCGAGACAGTCTTCTGAGCTTCTGAGTGTGCTTCAAGAGTGTGCTTCC 1265
DB 1209 GACTGAATCCGAGACAGTCTTCTGAGCTTCTGAGTGTGCTTCAAGAGTGTGCTTCC 1268
QY 1266 ACTGATCAAGCTCGATATACAGAGATCTGATGACCTGTCTTCACTTATGAGTTGG 1325
DB 1269 ACTGATCAAGCTCGATATACAGAGATCTGATGACCTGTCTTCACTTATGAGTTGG 1328
QY 1326 GACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTTGTCAAGCCGACAGCTGA 1385

Db 553 ACATGGGATCAGATGCTCCGGGAACTGGGCTTCAAGGACGAGTGGCTGTCTCC 612
 Qy 631 TGGGTCAGAAAGACATCGAGTCTTCGGTGGGACCCCGAGCTGTGACCATCTTGGC 690
 Db 613 TGGGTCAGAAAGACATCGAGTCTTCGGTGGGACCCCGAGCTGTGACCATCTTGGC 672
 Qy 691 GAGTCCGGGAGCCATTAAGTGTTCCTAGTCTTAATCTGTCTCCCATGGCCAAAGCTTA 750
 Db 673 GAGTCCGGGAGCCATTAAGTGTTCCTAGTCTTAATCTGTCTCCCATGGCCAAAGCTTA 732
 Qy 751 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCTCTTAAGGCGCCATGAT 810
 Db 733 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCTCTTAAGGCGCCATGAT 792
 Qy 811 TATGAGAAAGTGAAGACCTGAGAGTGTTCGACATTTCTGTTGGTAAACAATGGTCCAGC 870
 Db 793 TATGAGAAAGTGAAGACCTGAGAGTGTTCGACATTTCTGTTGGTAAACAATGGTCCAGC 852
 Qy 871 TCTGAGGCGCTGTGAGGTGCTGAGAACMAAACCCTCCAAAGAGCTGCTGACCTCAGC 930
 Db 853 TCTGAGGCGCTGTGAGGTGCTGAGAACMAAACCCTCCAAAGAGCTGCTGACCTCAGC 912
 Qy 931 CAGAAAACAAAGTCTTCACTCGAGTGTGATGAGTGTCTTCTTCTAATAGCTCTTA 990
 Db 913 CAGAAAACAAAGTCTTCACTCGAGTGTGATGAGTGTCTTCTTCTAATAGCTCTTA 972
 Qy 991 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATACCA 1050
 Db 973 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATACCA 1032
 Qy 1051 GAGTGTGCTTCTGCTGTGCTTAATGAAGAGGCTCCTGAGATCTCATGAGTCCAAAG 1110
 Db 1033 GAGTGTGCTTCTGCTGTGCTTAATGAAGATTCGGCTGTCCATATGCTCCCTCA 1092
 Qy 1111 TCCCTT-----GCCCTCATGATATCAAAACATCCTGACATC 1149
 Db 1093 AACCTGATGAGAGCTTTGGCTTCAACAGCTGGGCAATTCACAGAAAGACATLGAAGATC 1152
 Qy 1150 CCGCTCAGATTTGACCTTGTGCTTAATGAATATCTTCATGACAGACATCCTCTGATC 1209
 Db 1153 CCGCTCAGATTTGACCTTGTGCTTAATGAATATCTTCATGACAGACATCCTCTGATC 1212
 Qy 1210 GAAATCCGAGACAGTCTTGTGAAGACACGAGCGGCTTTGTCAAAAGCCGACGCTGATGA 1269
 Db 1213 GAAATCCGAGACAGTCTTGTGAAGACACGAGCGGCTTTGTCAAAAGCCGACGCTGATGA 1272
 Qy 1270 ATCAAGCTCAGATATCAGAGAGTCTGTGACACCTGTCTAATCTTATGAGTTCCGGAC 1329
 Db 1273 ATCAAGCTCAGATATCAGAGAGTCTGTGACACCTGTCTAATCTTATGAGTTCCGGAC 1332
 Qy 1330 CCGCTCAGTGTCTTGAAGACACGAGCGGCTTTGTCAAAAGCCGACGCTGATGA 1389
 Db 1333 CCGCTCAGTGTCTTGAAGACACGAGCGGCTTTGTCAAAAGCCGACGCTGATGA 1392
 Qy 1390 GTCCGCTTGTGTCCGAGTGTGCTTCTGAAGGGGAGCATTTGTATGTTGAAAGAGCC 1449
 Db 1393 GTCCGCTTGTGTCCGAGTGTGCTTCTGAAGGGGAGCATTTGTATGTTGAAAGAGCC 1452
 Qy 1450 ACGGAGAGAGAAATTACTGAGCCGAGAAAGATGAAATATCTGGGCTACCTTGTCTGA 1509
 Db 1453 ACGGAGAGAGAGAAATTACTGAGCCGAGAAAGATGAAATATCTGGGCTACCTTGTCTGA 1512
 Qy 1510 ACCGGAAATCTTAATGGGAAGACCTGTCTTGTGGCCAGCTTAATATCTGAGTGAAG 1569
 Db 1513 ACCGGAAATCTTAATGGGAAGACCTGTCTTGTGGCCAGCTTAATATCTGAGTGAAG 1572
 Qy 1570 TACCTCAGCTGAGCTTGAACATGAGCTCGGACAGAGACTCAAAAGAACCGCGGTGAT 1629
 Db 1573 TACCTCAGCTGAGCTTGAACATGAGCTCGGACAGAGACTCAAAAGAACCGCGGTGAT 1632
 Qy 1630 TTTTGG 1635
 Db 1633 GTGTGG 1638

RESULT 9
 ADR19681
 ID ADR19681 standard; DNA; 1857 BP.
 XX
 AC ADR19681;
 XX
 AC 07-OCT-2004 (first entry)
 DT
 XX
 DE Human drug metabolizing enzyme (DME)-2 gene sequence.
 XX
 KW drug metabolizing enzyme; DME; cytostatic; immunosuppressive;
 KW antiinflammatory; endocrine; ophthalmological; gastrointestinal;
 KW hepatocytic; cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; endocrine disorder; eye disorder;
 KW gastrointestinal disorder; liver disorder; metabolic disorder; gene; ds;
 KW human.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS 1. 1857
 FT /tag= a
 FT /product= "Human drug metabolizing enzyme (DME) 2"
 XX
 PN WO200226988-A2.
 PD
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US030662.
 XX
 PR 29-SEP-2000; 2000US-0236947P.
 PR 06-OCT-2000; 2000US-023864P.
 PR 20-OCT-2000; 2000US-0242323P.
 PR 09-NOV-2000; 2000US-0247581P.
 PR 16-NOV-2000; 2000US-0249519P.
 PR 22-NOV-2000; 2000US-0252834P.
 PR 30-NOV-2000; 2000US-0250567P.
 PR
 XX
 PA (INCY-.) INCYTE GENOMICS INC.
 XX
 PI Azimzai Y, Baughin MR, Borowsky ML, Ding L, Duggan BM;
 PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJB, Ison CH, Khan FA;
 PI Lal P, Lee EA, Lu DM, Nguyen DB, Arvizu C, Policky JU, Ramkumar J;
 PI Ring HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK;
 PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 XX
 DR WPI; 2002-362498/39.
 DR P-PSDB; ADR19663.
 XX
 PT Human drug metabolizing enzymes, useful in the diagnosis and treatment of
 PT disorders associated with aberrant (DME) activity, e.g., cancer and
 PT autoimmune disorders.
 PS
 PS Claim 12; SEQ ID NO 20; 142bp; English.
 XX
 CC This invention relates to novel drug metabolizing enzymes (DME) and the
 CC nucleotide sequences which encode them. The invention may be useful for
 CC the development of compounds with a cytostatic, immunosuppressive,
 CC antiinflammatory, endocrine, ophthalmological, gastrointestinal or
 CC hepatocytic activity acting as an agonist or antagonist of drug
 CC metabolizing enzyme activity. The invention may be used in the diagnosis
 CC and treatment of disorders associated with decreased or increased
 CC expression or activity of drug metabolizing enzymes. Such disorders
 CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
 CC endocrine, eye, gastrointestinal (including liver disorders) and
 CC metabolic disorders. The present sequence is that of a gene which encodes
 CC a human drug metabolizing enzyme (DME) of the invention. Note: This
 CC sequence did not form part of the printed specification but was obtained
 CC in electronic format from EPO.
 CC
 CC Sequence 1857 BP; 428 A; 487 C; 496 G; 446 T; 0 U; 0 Other;

Query Match 79.4%; Score 1386.2; DB 7; Length 1857;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

```

QY 81 GGGAAAACTGGGCTTCTGTCTGAAGGCGCAGAGAAACACGAGCTGGATGATTC 140
DB 264 GGTTCAGAGAGGCGCTTCTGTCTGAAGGCGCAGAGAAACACGAGCTGGATGATTC 323
QY 141 GGGCAAGCAAGTCACTGTCTGGAAAGCCCTGTGCTTGAACGTCTTCGAGATCCC 200
DB 324 GGGCAAGCAAGTCACTGTCTGGAAAGCCCTGTGCTTGAACGTCTTCGAGATCCC 383
QY 201 CTTTGTGTCCCGCGTGGAGTCCCTGGATTGAACCCGAGCTGCATCCGCTTG 260
DB 384 CTTTGTGTCCCGCGTGGAGTCCCTGGATTGAACCCGAGCTGCATCCGCTTG 443
QY 261 GGATTAACCTTGCAGAGAGCACTCTCAATTTGTGCTTCAGAACTCAGAGTGGCT 320
DB 444 GGATTAACCTTGCAGAGAGCACTCTCAATTTGTGCTTCAGAACTCAGAGTGGCT 503
QY 321 GCTTTAGATCAACATCTCTCAAGGTGCAATTAACCGAAATTCGAGTCTCAGAACTG 380
DB 504 GCTTTAGATCAACATCTCTCAAGGTGCAATTAACCGAAATTCGAGTCTCAGAACTG 563
QY 381 CCTTACCTGAACATCTTATGCGCGTCCGACCGCGATAGAGGCTCCAGCTCCGCTTT 440
DB 564 CCTTACCTGAACATCTTATGCGCGTCCGACCGCGATAGAGGCTCCAGCTCCGCTTT 623
QY 441 GGTGTGTCCAGAGAGTGTCTTCAAGACTGGCTCAGCTCCATCTTGTATGGTCCGC 500
DB 624 GGTGTGTCCAGAGAGTGTCTTCAAGACTGGCTCAGCTCCATCTTGTATGGTCCGC 683
QY 501 CCTGTGTCTTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
DB 684 CCTGTGTCTTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743
QY 561 TTTTCTTACCAATGAGGATCAGACATGCTCCGCGGAACTGGGCTTCAAGAACAGTGGC 620
DB 744 TTTTCTTACCAATGAGGATCAGACATGCTCCGCGGAACTGGGCTTCAAGAACAGTGGC 803
QY 621 TGTCTGTCTGTGGGTCTCAGAAAGACATCGAGTCTTCTGCTGGGAGCCCGAGCTGTGAC 680
DB 804 TGTCTGTCTGTGGGTCTCAGAAAGACATCGAGTCTTCTGCTGGGAGCCCGAGCTGTGAC 863
QY 681 CATCTTTGGCGAGTCCGCGGAGCCATAAGTGTCTTATCTGTCTTCCATGGC 740
DB 864 CATCTTTGGCGAGTCCGCGGAGCCATAAGTGTCTTATCTGTCTTCCATGGC 923
QY 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTAAGTGA 800
DB 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTAAGTGA 983
QY 801 GGCCCATGATTAAGAGAGTGAAGACCTGCAAGGTGTGTGACATTTCTGTGTAAACA 860
DB 984 GGCCCATGATTAAGAGAGTGAAGACCTGCAAGGTGTGTGACATTTCTGTGTAAACA 1043
QY 861 TGGGTCAAGCTTGAAGCCCTGTGAGGTGCTTGAAGCAAAACCTTCCAGAGTGTCT 920
DB 1044 TGGGTCAAGCTTGAAGCCCTGTGAGGTGCTTGAAGCAAAACCTTCCAGAGTGTCT 1103
QY 921 GACCTTCAGCCGAGAAACAAAGTCTTTCATCTGAGTGTGTATGTGTGTCTTCTTAA 980
DB 1104 GACCTTCAGCCGAGAAACAAAGTCTTTCATCTGAGTGTGTATGTGTGTCTTCTTAA 1163
QY 981 TGAAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT 1040
DB 1164 TGAAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT 1223
QY 1041 CAATAACCAAGAGTGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1100
DB 1224 CAATAACCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1255

```

```

QY 1101 CTCCAACAAGTCCCTTGTCCCTCCATCTGATGATACAAAAATCTGTGACATCCGCTCAGTA 1160
DB 1256 -----TGCATATCCGCTCAGTA 1274
QY 1161 TTTCGACTTGTGGCTTAATGAATACTTCATGACAAAGCATCCCTGACTGAAATCCGAGA 1220
DB 1275 TTTCGACTTGTGGCTTAATGAATACTTCATGACAAAGCATCCCTGACTGAAATCCGAGA 1334
QY 1221 CAGTCTTCTGACCTTGTGAGATGTGTCTTTGTGTGCTCCCTGACATGATCAGACTGG 1280
DB 1335 CAGTCTTCTGACCTTGTGAGATGTGTCTTTGTGTGCTCCCTGACATGATCAGACTGG 1394
QY 1281 ATATCACAAGATGTCTGTGACACTGTCTACTTATGAGTTTGGCACCAGCTCAGTG 1340
DB 1395 ATATCACAAGATGTCTGTGACACTGTCTACTTATGAGTTTGGCACCAGCTCAGTG 1454
QY 1341 CTTTGAAAGCAGAAAGCCGCTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTTGT 1400
DB 1455 CTTTGAAAGCAGAAAGCCGCTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTTGT 1514
QY 1401 GTTTCGTGTGCTTCTTGAAGGGGACATTTGTATGTTCGAAGAGCCACGAGAGAGA 1460
DB 1515 GTTTCGTGTGCTTCTTGAAGGGGACATTTGTATGTTCGAAGAGCCACGAGAGAGA 1574
QY 1461 GAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTGTCTGAAACGGGAATCC 1520
DB 1575 GAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTGTCTGAAACGGGAATCC 1634
QY 1521 TAATGGAAAGCATGTGTCTGTGTGCGCAGCTTATATCTGACTGAGACGTACTCGAGT 1580
DB 1635 TAATGGAAAGCATGTGTCTGTGTGCGCAGCTTATATCTGACTGAGACGTACTCGAGT 1694
QY 1581 GGACTTTGAACATGAGCTTGTGACAGAGACTCAAGAAACCGCGGGTGAATTTTGG 1635
DB 1695 GGACTTTGAACATGAGCTTGTGACAGAGACTCAAGAAACCGCGGGTGAATTTTGG 1749

RESULT 10
ADA53168
ID ADA53168 standard; cDNA; 1962 BP.
XX
AC ADA53168;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human coding sequence, SEQ ID 736.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
  Gene Therapy; human; secretory protein; membrane protein; cancer;
  inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SBP-2001; 2001JP-00328381.
XX
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
  Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
  Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
  WPI; 2003-395539/38.
DR P-PSDB; ADA54807.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory

```

PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 PS Claim 1; SEQ ID NO 736; 205bp; English.

CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA5433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 1962 BP; 428 A; 541 C; 473 G; 520 T; 0 U; 0 Other;

Query Match 77.0%; Score 1344.4; DB 10; Length 1962;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

QY 91 GGGGCTTCTGCTGAAGGCGCAGAGAGAACACAGGCTGGATGATTACAGGCGAAGCAA 150
 DB 222 GGGCTTCTGCTGAAGGCGCAGAGAGAACACAGGCTGGATGATTACAGGCGAAGCAA 281
 QY 151 GTCACTGTCTGGAGAGCCCTGTGCTGTGAAGCTGTCTGGAGTCCCTTTGCTGCT 210
 DB 282 GTCACTGTCTGGAGAGCCCTGTGCTGTGAAGCTGTCTGGAGTCCCTTTGCTGCT 341
 QY 211 CCCCCGCTGGAGTCCCTGTGCTGTGAAGCCGAGACCTGCATCGCCCTGGATTAATTG 270
 DB 342 CCCCCGCTGGAGTCCCTGTGCTGTGAAGCCGAGACCTGCATCGCCCTGGATTAATTG 401
 QY 271 CGAGAGACACCTCCACCTTAATTGTCCTCCAGAACTCAGAGTGGTCTCTTAAGT 330
 DB 402 CGAGAGACACCTCCACCTTAATTGTCCTCCAGAACTCAGAGTGGTCTCTTAAGT 461
 QY 331 CAACAATGCTCAAGGTGCTATACCGAAATTCGAGTGTCAAGAACTGCTCTACCG 390
 DB 462 CAACAATGCTCAAGGTGCTATACCGAAATTCGAGTGTCAAGAACTGCTCTACCG 521
 QY 391 AACATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 450
 DB 522 AACATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
 QY 451 CCAGAGAGTGCCTTCAAGACTGCTCAGGCTCATCTTGAATGGTCCGCGCTGCTGCT 510
 DB 582 CCAGAGAGTGCCTTCAAGACTGCTCAGGCTCATCTTGAATGGTCCGCGCTGCTGCT 641
 QY 511 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
 DB 642 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
 QY 571 ACATGGAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
 DB 702 ACATGGAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 761
 QY 631 TGGGTCCAGAGAACTCAGAGTCTTCCGATGGGAGCCCACTGCTGTAACATCTTGGC 690
 DB 762 TGGGTCCAGAGAACTCAGAGTCTTCCGATGGGAGCCCACTGCTGTAACATCTTGGC 821
 QY 691 GAGTCCCGGGAGCCATTAAGTCTTCTTAAGTCTTCTTCTCCATGGCCAAAGCTTA 750
 DB 822 GAGTCCCGGGAGCCATTAAGTCTTCTTAAGTCTTCTTCTCCATGGCCAAAGCTTA 881
 QY 751 TTCCACAAGGCTCATGAGAGAGTGGGTGGCATATCCCTTACCTGAGAGCCCAATGAT 810
 DB 882 TTCCACAAGGCTCATGAGAGAGTGGGTGGCATATCCCTTACCTGAGAGCCCAATGAT 941
 QY 811 TATGAGAAAGTGAAGACTGCAAGTGTGTCATCTTCTGTTAACTGCTGACAG 870
 DB 942 TATGAGAAAGTGAAGACTGCAAGTGTGTCATCTTCTGTTAACTGCTGACAG 1001
 QY 871 TCTGAGGCTGCTGAGAGTGTCTGAGAGCAAAACCTTCCAGAGGCTGAGCCCTGAGC 930
 DB 1002 TCTGAGGCTGCTGAGAGTGTCTGAGAGCAAAACCTTCCAGAGGCTGAGCCCTGAGC 1061

QY 931 CAGAAAACAAAGCTTTTCACTGAGTGGTGAATGGTCTTTCTTCTTAATGAGCTCTTA 990
 DB 1062 CAGAAAACAAAGCTTTTCACTGAGTGGTGAATGGTCTTTCTTCTTAATGAGCTCTTA 1121
 QY 991 GATCTATTGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGAGAGTCAATAACAC 1050
 DB 1122 GATCTATTGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGAGAGTCAATAACAC 1181
 QY 1051 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1110
 DB 1182 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1241
 QY 1111 TCCCTTGCCCTTCATCTGATCAAAACATCTGCAATCCCGCTCATATTTGACCTT 1170
 DB 1242 TCCCTTGCCCTTCATCTGATCAAAACATCTGCAATCCCGCTCATATTTGACCTT 1301
 QY 1171 GTGGCTAATGAATCTTCATGACAGAGACTCCCTGATCTGAATCCGAGACAGTCTTGTG 1230
 DB 1302 GTGGCTAATGAATCTTCATGACAGAGACTCCCTGATCTGAATCCGAGACAGTCTTGTG 1361
 QY 1231 GACTTGTGAGAGATGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290
 DB 1362 GACTTGTGAGAGATGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421
 QY 1291 GATGCTGAGTCACTGCTCTAATTATGAGTTCGAGACCGGCTCAAGTCTTTGAAGAC 1350
 DB 1422 ----- 1421
 QY 1351 AGGAAGCGGCTTTTGTCAAGCCGACACGCTGATGATGCTCGCTTGTGCTGCTGCT 1410
 DB 1422 ----- 1421
 QY 1411 GCCTTCTGAAGGGGGACATTTGTTATGTTGGAAGAGACCCAGAGAGAGAAATTACTG 1470
 DB 1422 ----- 1451
 QY 1471 AGCCGAGAGATGATGAATATCTGGCTACTTGTGCTGGAACCGGGAATCTTAATGGGAC 1530
 DB 1452 AGCCGAGAGATGATGAATATCTGGCTACTTGTGCTGGAACCGGGAATCTTAATGGGAC 1511
 QY 1531 GACTGTCTCTGTGGCCAGCTTATATCTGATGAGAGTACCTCCAGCTGGAATTGAAC 1590
 DB 1512 GACTGTCTCTGTGGCCAGCTTATATCTGATGAGAGTACCTCCAGCTGGAATTGAAC 1571
 QY 1591 ATGAGCTCGGAGAGAGACTCAAAAGAACCGGGGTGATTTTGGACCAAGACATATCC 1650
 DB 1572 ATGAGCTCGGAGAGAGACTCAAAAGAACCGGGGTGATTTTGGACCAAGACATATCC 1631
 QY 1651 CTGATCTGTCTGCTCCGACATGCTCCACAGTCTCTTCTTCTTAATCTTCTCTCT 1710
 DB 1632 CTGATCTGTCTGCTCCGACATGCTCCACAGTCTCTTCTTCTTAATCTTCTCTCT 1691
 QY 1711 CTCCTCAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTGA 1746
 DB 1692 CTCCTCAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTGA 1727

RESULT 11
 ADF50146
 ID ADF50146 standard; cDNA; 2145 BP.
 XX
 AC ADF50146;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Cat cauxin encoding cDNA SEQ ID NO:3.
 XX
 KW cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
 XX
 OS Felis catus.
 XX
 PN JP2003250575-A.
 XX

PD 09-SEP-2003.
XX 04-MAR-2002; 2002JP-00057908.
XX 04-MAR-2002; 2002JP-00057908.
XX (TOHO-) TOHOKU TECHNOARCH KK.
XX MPI: 2004-002277/01.
XX P-PSDB; ADP50147.
XX Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
XX
XX Claim 6; SEQ ID NO 3; 33pp; Japanese.
XX
XX The present sequence encodes a cat cauxin protein (I) or its salt, which
XX is cat kidney disease marker. Also described: (1) a partial peptide (II)
XX of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
XX comprising (III); (4) a transformed host (V) comprising (III) or (IV);
XX (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
XX couples specifically with (I) or (II); (7) diagnosing cat kidney disease
XX which involves measuring (I) quantitatively, and where reduction of
XX amount of (I) indicates presence of the disease; (8) a cat kidney disease
XX diagnostic agent comprising (I) labelling agent, a reagent which measures
XX the biological activity of urinary (I) or (VI); and (9) a cauxin
XX detection kit which measures cauxin in a test sample. (I) is useful as a
XX cat kidney disease marker and (VI) is useful for diagnosing cat kidney
XX disease. (I) enables detection of cat kidney disease simply and
XX correctly. (I) provides an early marker for the disease, and replaces
XX complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
XX blood testing.

XX Sequence 2145 BP; 444 A; 581 C; 550 G; 564 T; 0 U; 6 Other;

Query Match 62.9%; Score 1099; DB 12; Length 2145;
Best Local Similarity 80.1%; Pred. No. 8.6e-314;
Matches 1329; Conservative 0; Mismatches 323; Indels 8; Gaps 3;

QY 91 GGGCTTGTCTGTAAGGGCCACAGAGAACACACGCTGGATGATTCAGGGCAAGCAA 150
DB 238 GGGCCAGCTGCTGAGAGCAACAGTGAAGAGCACAGGCTGGATGGTCCGGGAAAGCAA 297
QY 151 GTCACTGCTGGGAAGCCGTGCTGTGAACGTTCTTCGAGTCCCTTGTCTGT 210
DB 298 ACGACTGACTGGGAAGCACGCTGCTGTGAACGTTCTTCGAGTCCCTTGTCTGT 357
QY 211 CCCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCATCCCTGGGATACTTG 270
DB 358 CTTCTCTAGGGCCCTCGTGAATTAAGCAACAAAGCCGTGCTGCTCCGGGAATGACTTC 417
QY 271 CGAGAAAGCACTCTCACTTAATTTGTGCTCTCAAGAACTCAAGATGCTGCTTTAAT 330
DB 418 CGAATATGCAATCTTACCTTAATTAATGCTTCAAGAACTTGAATGCTGCTTCTTAAT 477
QY 331 CAACATGCTCAAGTGCATTAACCGAAATTCGAGATGCAAGATGCTGCTTCACTG 390
DB 478 CAACATGCTCAAGTGCATTAACCGAAATTCGAGATGCTGCTGCTTCACTG 537
QY 391 AACATCTATGCGCTGCCACGCGCATACAGGCTCCAGCTCCCGCTTGTGTTGTTT 450
DB 538 AACATCTATGCGCGACGCCCATGCGCAATGCTCCAGCTCCCGCTTGTGTTGTTT 597
QY 451 CGAGAGGCTGCTTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
DB 598 CCGCGGGGCTGCTTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
QY 511 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 658 TACGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 571 ACATGAGATCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630

DB 718 ACAGGGAGTGAAGATGCCCGGGGAACTGGGCTTGTGAGACAGATGCTGCCCTTAAC 777
QY 631 TGGGTCCAGGAAGATGAGTTCTTGGTGGGGACCCAGCTGTGATGCATCTTTGGC 690
DB 778 TGGGTCCGGGAAACATGAGTTCTTGGTGGGGACCCAGCTGTGATGCATCTTTGGC 837
QY 691 GAGTCCGGGGAGCCATGAGTTCTTGGTGGGGACCCAGCTGTGATGCATCTTTGGC 750
DB 838 GAGTCCGGGGAGCCATGAGTTCTTGGTGGGGACCCAGCTGTGATGCATCTTTGGC 897
QY 751 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACCTTGAG--GCCAT 807
DB 898 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACCTTGAGAGCCCT 957
QY 808 GATTATGAGAGAGTGAAGACCTGAGAGTGGTGAACATTTCTGATGAACAATGCTCA 867
DB 958 GATTATGAGAGAGAGAGATTTGAGAGTGGTGGCCATCATCTTACCTTGAGAGCT 1017
QY 868 GACTCTGAGGCGCTCTGAGTGGCTGAGACAAACCTCCAGAGAGCTGCTGACCCCT 927
DB 1018 GACTCTGAGGCGCTCTGAGTGGCTGAGACAAACCTCCAGAGAGTGAATGAGACATC 1077
QY 928 AGCCAGAAACAAAGTCTTTCACTGAGTGGTGAATGAGTGGCTTTCTTAATAGCT 987
DB 1078 AGCAAGAAACTCAAGTTTTCATTCAGATGATGATGATTTCTTCTGATAGCT 1137
QY 988 CTAGATCTATGCTCAAGAAAGATTTAAGCAATTCCTTCATCATGAGATCAATAC 1047
DB 1138 GTAGGCTATGATCAAAAGAGATTTAATGATTTCTTATATGAGATCAATAC 1197
QY 1048 CACGAGTGGCTTCTGCTGCTGATGAGAGGCTCTGAGATCTGAGTGGCTCAAC 1107
DB 1198 CACGAGTGGCTTCTGCTGCTGATGAGAGGCTCTGAGATCTGAGTGGCTCAAC 1254
QY 1108 AAGTCCCTTGGCTTCATCTGATCAAAACATCTGCAATCCGCTCAATATTGCA 1167
DB 1255 AAGTCTGAGCCCTCTACTAGTACACAGTTCTGAAATATTCACCAAGATTTGCA 1314
QY 1168 CTGTGCTATGATTAATTTCTCATGACAAAGCATCTCCGATGAAATCCGAGACGTCT 1227
DB 1315 CTGTGCTATGATTAATTTCTCATGACAAAGCATCTCCGATGAAATCCGAGATGTTT 1374
QY 1228 CTGAGCTTGTGAGATGTTCTTGTGCTGCTGATGATGATGATGATGATGATGAT 1287
DB 1375 CTGAGCTTGTGAGATGTTCTTGTGCTGCTGATGATGATGATGATGATGATGAT 1434
QY 1288 AGAGATGCTGAGCACTGTCTACTTATGAGTGTGGCAACCGGCTCACTGCTTTGA 1347
DB 1435 AGAGATGCTGAGCACTGTCTACTTATGAGTGTGGCAACCGGCTCACTGCTTTGA 1494
QY 1348 GACAGAAAGCGGCTTTTGTCAAAGCGACACAGCTGATGAAAGTCCGCTTGTGGT 1407
DB 1495 GACAGAAAGCGGCTTTTGTCAAAGCGACACAGCTGATGAAAGTCCGCTTGTGGT 1554
QY 1408 GGTGCTTCTGAAAGGGGCAATTTGTTGAAAGAGCCAGAGAGAGAGAAAGTTA 1467
DB 1555 GGTGCTTCTGAAAGGGGCAATTTGTTGAAAGAGCCAGAGAGAGAGAAAGTTG 1614
QY 1468 CTGAGCCGGAAGATGATGAATACTGAGCTACCTTTGCTGAACCGGAAATCTTAAGG 1527
DB 1615 CTGAGCAAGAAAGATGATGAGTATGAGGCTACCTTTGCTGAGCGGGAACCTTAAGG 1674
QY 1528 AAGCACTGTCTGTGCGCAGCTTATATCTGACTGAGCAATCTCACTGAGACTTG 1587
DB 1675 GAAAGTGTGCTGTGTGCGCAGCTTACCAAGAGCGAGATCACTTAAGTGAATTTG 1734
QY 1588 AAGCACTGTCTGTGCGCAGCTTATATCTGACTGAGCAATCTCACTGAGACTTG 1646
DB 1735 AAGTGTGAGCTGTGAGCAAGAACTGAAGAGCAAGAGTGTGAGTGAATGCAAT 1794
QY 1647 -CCCCCTGATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCACTTACCTTCC 1705
DB 1795 GTCCCTGATACCCCCCACTCCAGAGGCTTCCCAAGTCTCTTCTTCACTTACCTTCC 1854

QY 1706 TCTCTCTCCACGCTTTCTTTCTTTCTGCTCTTG 1745
DB 1855 TTCTTTGCTCCGCTGCTTTTCTGCTCATG 1894

RESULT 12
ID ADF50144 standard; cDNA; 1629 BP.
AC ADF50144;

DT 12-FEB-2004 (first entry)

DE Cat cauxin encoding cDNA SEQ ID NO:1.

KM cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.

OS Feline cat.

PN JP2003250575-A.

PD 09-SEP-2003.

PF 04-MAR-2002; 2002JP-00057908.

PR 04-MAR-2002; 2002JP-00057908.

PA (TOHO-) TOHOKU TECHNARCH KK.

DR WPI; 2004-002277/01.

DR P-PSDB; ADF50145.

PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
and for diagnosing cat kidney disease.

XX Claim 6; SEQ ID NO 1; 33bp; Japanese.

XX The present sequence encodes a cat cauxin protein (I) or its salt, which
is cat kidney disease marker. Also described: (1) a partial peptide (II)
of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
comprising (III); (4) a transformed host (V) comprising (III) or (IV);
(5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
couple specifically with (I) or (II); (7) diagnosing cat kidney disease
which involves measuring (I) quantitatively, and where reduction of
amount of (I) indicates presence of the disease; (8) a cat kidney disease
diagnostic agent comprising (I) labelling agent; a reagent which measures
the biological activity of urinary (I) or (VI); and (9) a cauxin
detection kit which measures cauxin in a test sample. (I) is useful as a
cat kidney disease marker and (VI) is useful for diagnosing cat kidney
disease. (I) enables detection of cat kidney disease simply and
correctly. (I) provides an early marker for the disease, and replaces
complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
blood testing.

XX Sequence 1629 BP; 345 A; 438 C; 441 G; 405 T; 0 U; 0 Other;

Query Match 60.9%; Score 1063.2; DB 12; Length 1629;
Best Local Similarity 80.9%; Pred. No. 3e-303;
Matches 1265; Conservative 0; Mismatches 293; Indels 6; Gaps 2;

QY 91 GGGGCTTCTGTGAAGGCCACAGAGAACACAGGCTGGAGATTCAAGGGCAACAA 150
DB 64 GGGGCACTGTGTGAAGGCCACAGAGAACACAGGCTGGAGATTCAAGGGCAACAA 123
QY 151 GTCACTGTGTGAAGGCCACAGAGAACACAGGCTGGAGATTCAAGGGCAACAA 127
DB 124 ACACCTTACTGTGAAGGCCACAGAGAACACAGGCTGGAGATTCAAGGGCAACAA 183
QY 211 CCCCCGTGGAGATCCCTGCGATTTCAGAACCCGAGCCTGCATGCGCTGGAGATTCAAG 270
DB 184 CTTCTCTTAAGGGCCCTGCGATTTCAGAACCCGAGCCTGCATGCGCTGGAGATTCAAG 243

QY 271 CGAAGACCACTCTTACCTTAATTTGTGCTCCAGAACTGAGATGCTGCTTAAGT 330
DB 244 CGAAATGCCAATCTTACCTTAATTTGTGCTCCAGAACTGAGATGCTGCTTAAGT 303
QY 331 CAACATGCTCAAGTGTATTAACCCGAATTTGAGATGTCAAGAACTGCTTACTG 390
DB 304 CAACAGTTCTCAAGTGTATTAACCCGAATTTGAGAAAGTGTCAAGAACTGCTTACTG 363
QY 391 AACATCTATGAGCGCTGCCACGCGCATACAGAGCTCAAGCTCCCGTGTGTGTGCTC 450
DB 364 AACATCTATGAGCGCGCTGCCACGCGCATACAGAGCTCAAGCTCCCGTGTGTGTGCTC 423
QY 451 CCAGAGGTGCTTCAAGATGAGCTGAGCTTCAAGCTTCAAGTGTGTGTGTGTGTGTGTG 510
DB 424 CCGGCGGTGCTTCAAGATGAGCTGAGCTTCAAGCTTCAAGTGTGTGTGTGTGTGTGTG 483
QY 511 TATGAGACGCTG 570
DB 484 TACGAGACGCTG 543
QY 571 ACATGAGATCAGATGCTCCGCGGAATGAGCTTCAAGACAGAGTGTGTGTGTGTGTGTG 630
DB 544 ACAGGAGATGAGATGCTCCGCGGAATGAGCTTCAAGACAGAGTGTGTGTGTGTGTGTG 603
QY 631 TGGGTCCAGAAAGAACTGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 690
DB 604 TGGGTCCGAGCAACATGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 663
QY 691 GAGTCCGCGGAGCATATGATGTTTCTAGTCTTAATCTGTCTCCAGAGCAAGGCTTA 750
DB 664 GAGTCCGCGGAGCATATGATGTTTCTAGTCTTAATCTGTCTCCAGAGCAAGGCTTA 723
QY 751 TTCCAGAAAGCATATGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 807
DB 724 TTCCAGAAAGCATATGAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 783
QY 808 GATTATGAGAAAGTG 867
DB 784 GGTATATGAGAAAGTG 843
QY 868 GACTGTGAGGCTG 927
DB 844 GACTGTGAGGCTG 903
QY 928 AGCCAGAAAGAAAGTG 987
DB 904 AGCCAGAAAGAAAGTG 963
QY 988 CTAGATCTATTTGTCTAGAAAGATTTAAAGCAATTTCTTCAATCTGAGATCAATAC 1047
DB 964 GTAGCCCTATTTGTCTAGAAAGATTTAAAGCAATTTCTTCAATCTGAGATCAATAC 1023
QY 1048 CACGAGTGTGCTTG 1107
DB 1024 CACGAGTGTGCTTG 1080
QY 1108 AAGTCCCTGTGCTTG 1167
DB 1081 AAGTCCCTGTGCTTG 1140
QY 1168 CTTGTGGCTAATGAATCTTCAAGACAGCACTCCCTCAAGTGAATCCGAGAGCTGT 1227
DB 1141 CTTGTGGCTAATGAATCTTCAAGACAGCACTCCCTCAAGTGAATCCGAGAGCTGT 1200
QY 1228 CTGGAATG 1287
DB 1201 CTGGAATG 1260
QY 1288 AGAGATG 1347
DB 1261 AGAGATG 1320
QY 1348 GACAGAAAGCCGCTTTGTCTAAAGCCGACAGCTGTATGAATCCGCTTTGTGTGTGT 1407

Db 1321 GACACGAGGCCAGCTTTCGTGAAGCCGATCACTGATGAATCCGCTTCTTTGGA 1380
 Qy 1408 GGTGCTTCTCGAAGGGGACATTTGTATGTCGAAGAGCCACGAGAGAGAGAAATTA 1467
 Db 1381 GGTGCTTCTCGAAGGGGACATTTGTATGTCGAAGAGCCACGAGAGAGAGAAATTA 1440
 Qy 1468 CTGAGCCGGAAGATGATGAATACTAGGCTACTTGTCTGACCGGGAAATCTTAATGGG 1527
 Db 1441 CTGAGCAGGAAGATGATGAATACTAGGCTACTTGTCTGACCGGGAAATCTTAATGGG 1500
 Qy 1528 AACGACCTGTCTCTGTGGCCAGCTTATATCTGACTGAGCAGTACTCCAGCTGAGCTTG 1587
 Db 1501 GAAGGTGTGCTCTGTGGCCAGCTTATATCTGACTGAGCAGTACTCCAGCTGAGCTTG 1560
 Qy 1588 AACATGACCTCTGACAGAGATCTCAAGAAACCGCGGTGGATTTTGGACACGACCATTC 1647
 Db 1561 AGTGTGACCTGTGGGACAGAAATCTGAAGAGCAGAGGTGAGATTGGATATACCAT 1620
 Qy 1648 CCCC 1651
 Db 1621 GTCC 1624

RESULT 13

ABQ86171 standard: DNA; 1071 BP.

ABQ86171;

10-SEP-2002 (first entry)

Novel human gene. SEQ ID 42.

Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;
 neurotropic; neuroprotective; immunosuppressive; haemostatic;
 antiinflammatory; cardiant; antilucer; virucide; antithyroid;
 cerebroprotective; anorectic; metabolic; vaccinia; cancer; infection;
 wound healing disorders; atherosclerosis; Parkinson's disease;
 Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 inflammation; neoplastic disease; nervous system disorder;
 cardiovascular disorders; pancreatitis; respiratory disorder;
 hyperproliferation; systemic autoimmune disease; hyper-immunity;
 developmental abnormality; gastrointestinal ulceration; neuropathy;
 haematological disease; metabolic disease; sperm dysfunction;
 thyroid disorder; hypothyroidism; brain damage; colitis;
 cone photo- transduction deficiency; neurological disease; stroke;
 angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 trachea; thymus; lymph node; muscular system; obesity; anorexia;
 growth abnormality; precocious puberty; gene; ss.

Homo sapiens.

WO200250105-A1.

27-JUN-2002.

17-DEC-2001; 2001WO-US049232.

19-DEC-2000; 2000US-0256710P.
 20-DEC-2000; 2000US-0257048P.
 09-JAN-2001; 2001US-0260482P.
 30-JAN-2001; 2001US-0264922P.
 06-FEB-2001; 2001US-0266797P.
 19-MAR-2001; 2001US-0276988P.
 04-APR-2001; 2001US-0281535P.
 08-MAY-2001; 2001US-0289622P.

(SMIK) SMITHKLINE BEECHAM CORP.
 (SMIK) SMITHKLINE BEECHAM PLC.
 (GLAX) GLAXO GROUP LTD.
 Agrawal P, Birkeland M, Cogswell JP, Kahnlick KF, Lai Y;

PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
 XX MPI: 2002-508784/54.
 DR P-PSDB; ABP61006.
 XX
 PT Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX
 PS Claim 2(a); Page 252; 335pp; English.

The invention relates to an isolated polypeptide with signal sequences
 which allow it to be secreted extracellularly or membrane associated. The
 activity of polypeptides of the invention may be described as,
 CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, neurotropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiant, antilucer, virucide, antithyroid, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder;
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, angiogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
 CC of the invention

Sequence 1071 BP; 242 A; 286 C; 254 G; 289 T; 0 U; 0 Other;

Query Match 58.5%; Score 1021.4; DB 6; Length 1071;

Best Local Similarity 99.9%; Pred. No. 5-5e-291;

Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 724 ATACTGTCTCCAGTGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGCC 783
 Db 49 ATACTGTCTCCAGTGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGCC 108
 Qy 784 ATCATCCCTTACCTGAGAGCCCAATGATATGAGAAAGATGAGAGCTCAGAGTGTGCA 843
 Db 109 ATCATCCCTTACCTGAGAGCCCAATGATATGAGAAAGATGAGAGCTCAGAGTGTGCA 168
 Qy 844 CATTTCTGTGATTAACATGCGTCAGACTCTGAGGCGCTGTGAGTGCTGTGAGCAAAA 903
 Db 169 CATTTCTGTGATTAACATGCGTCAGACTCTGAGGCGCTGTGAGTGCTGTGAGCAAAA 228
 Qy 904 CCTTCGAAGAGCTGCTGAGCCCTCAGCCAGAAAACAAAGTCTTTCACTGAGTGTGAT 963
 Db 229 CCTTCGAAGAGCTGCTGAGCCCTCAGCCAGAAAACAAAGTCTTTCACTGAGTGTGAT 288
 Qy 964 GGTGCTTCTTCTTCAATGAGCCTTATGATCTATGTCTCAGAAAGCATTTAAACAAAT 1023
 Db 289 GGTGCTTCTTCTTCAATGAGCCTTATGATCTATGTCTCAGAAAGCATTTAAACAAAT 348
 Qy 1024 CCTTCATCATCGAGATCAATTAACAGAGTGTGCTTCGAGTCTCATGAAAGAGGCT 1083
 Db 349 CCTTCATCATCGAGATCAATTAACAGAGTGTGCTTCGAGTCTCATGAAAGAGGCT 408
 Qy 1084 CCTGAGATCTCAGTGGCTCCAAAGTCCCTTGCCCTCACTTGATCAAAAACATCTGTG 1143
 Db 409 CCTGAGATCTCAGTGGCTCCAAAGTCCCTTGCCCTCACTTGATCAAAAACATCTGTG 468
 Qy 1144 CACATCCGCGCTCAGATATTTGCACTTGTGGCTAATGAATCTTCCATGAGCAAGACATCC 1203
 Db 469 CACATCCGCGCTCAGATATTTGCACTTGTGGCTAATGAATCTTCCATGAGCAAGACATCC 528
 Qy 1204 CTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGAGATGATGTTCTTTGTGTCTCCT 1263

```
Db 529 CTGACGAAATCCGAGCAGCTCTTGGACCTTGCTGGAGATGTTCTTTGGGCTCT 588
QY 1264 GCATCATGACACTGATATCAAGAGATGCTGGTACCTGTCTATCTTATGAGTTT 1323
Db 589 GCATCATGACACTGATATCAAGAGATGCTGGTACCTGTCTATCTTATGAGTTT 648
QY 1324 CGGACACCGGCTCAGTGTCTTGAAGACAGAGCCGGCTTTTCTCAAGCCGACCGCT 1383
Db 649 CGGACACCGGCTCAGTGTCTTGAAGACAGAGCCGGCTTTTCTCAAGCCGACCGCT 708
QY 1384 GATGAAGTCGGCTTGTGTGCTGGTGTGCTTCTGAAGGGGACATGTTATGTTGAA 1443
Db 709 GATGAAGTCGGCTTGTGTGCTGGTGTGCTTCTGAAGGGGACATGTTATGTTGAA 768
QY 1444 GGAGCCAGGAGGAGAGAGTTACTGAGCCGGAGATATGAAATATCTGGCTTACTTT 1503
Db 769 GGAGCCAGGAGGAGAGAGTTACTGAGCCGGAGATATGAAATATCTGGCTTACTTT 828
QY 1504 GCTCGAACCAGGAGATCTTATGGAACGACCTGTCTGTGGCCAGCTTATATCTGACT 1563
Db 829 GCTCGAACCAGGAGATCTTATGGAACGACCTGTCTGTGGCCAGCTTATATCTGACT 888
QY 1564 GAGCAGTACTCTCAGTGTGATCTTGAACATGAGCTTGGACAGAGACTCAAGACCGCG 1623
Db 889 GAGCAGTACTCTCAGTGTGATCTTGAACATGAGCTTGGACAGAGACTCAAGACCGCG 948
QY 1624 GTGATTTTTGAGACGACATCCCTGATCTGTGCTGCTCGACATGCTCCACAGT 1683
Db 949 GTGATTTTTGAGACGACATCCCTGATCTGTGCTGCTCGACATGCTCCACAGT 1008
QY 1684 CCTCTTCTCTCTTAACTTCTCTCTCTCTCTCTCAAGCCCTTCTTCTTCTGCTCT 1743
Db 1009 CCTCTTCTCTCTTAACTTCTCTCTCTCTCTCTCAAGCCCTTCTTCTTCTGCTCT 1068
QY 1744 TGA 1746
Db 1069 TGA 1071

RESULT 14
ADCS5523
ID ADCS5523 standard; cDNA, 1244 BP.
XX
AC ADCS5523;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human carboxylase 24.64 encoding sequence.
XX
KW human carboxylase-24.64; primary hypertension; digestive ulcer;
KM nephrotic; bronchial asthma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 236..910
FT /*Cag= a
XX
PN CN1382799-A.
XX
PD 04-DEC-2002.
XX
PF 26-APR-2001; 2001CN-00112736.
XX
PR 26-APR-2001; 2001CN-00112736.
XX
PA (BIOW-) BIOWINDOM GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2003-269506/27.
XX
DR P-PSDB; ADCS5524.
DR
```

```
XX Polypeptide-human carboxylase-24.64 and polynucleotide for coding it.
PT
XX
XX Claim 6; SEQ ID NO 1; 31p; Chinese.
XX
XX The present invention relates to a polypeptide-human carboxylase-24.64,
CC the polynucleotide for coding it, the process for preparing the
CC polynucleotide by DNA recombination, the application of the polypeptide in
CC treating diseases such as primary hypertension, digestive ulcer,
CC nephrotic, bronchial asthma, tremor, etc, the antagonist of the polypeptide
CC and its medical action, and the application of the polynucleotide are
CC new. The present sequence represents human carboxylase 24.64 encoding
CC sequence.
XX
SQ Sequence 1244 BP; 301 A; 314 C; 271 G; 358 T; 0 U; 0 Other;
Query Match 51.8%; Score 905.2; DB 10; Length 1244;
Best Local Similarity 99.7%; Pred. No. 1.4e-256;
Matches 907; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 837 GGTTCACATTTCTGTGTAACTAATGCTGACCTTGAAGCCCTGTGAGTCTGAG 896
Db 1 GTTGCACATTTCTGTGTAACTAATGCTGACCTTGAAGCCCTGTGAGTCTGAG 60
QY 897 GACAAAACCTCCAAAGAGCTGACCTCAGCCAGCAAAAACAAAGTCTTCACTGAGT 956
Db 61 GACAAAACCTCCAAAGAGCTGACCTCAGCCAGCAAAAACAAAGTCTTCACTGAGT 120
QY 957 GGTGATGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1016
Db 121 GGTGATGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 180
QY 1017 AGCAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1076
Db 181 AGCAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 240
QY 1077 GAGGCTCTGAGATCTCTCAGTGGCTCAACAAGTCCCTGCTCATGATGATGATGATG 1136
Db 241 GAGGCTCTGAGATCTCTCAGTGGCTCAACAAGTCCCTGCTCATGATGATGATGATG 300
QY 1137 CATCTGACATCCCGCTCAGTATTTGACCTTGTGGCTAATGAAATCTTCCATGACAA 1196
Db 301 CATCTGACATCCCGCTCAGTATTTGACCTTGTGGCTAATGAAATCTTCCATGACAA 360
QY 1197 GCACTCCCTGACGAAATCCGAGACAGTCTTCTGAGCTTCTGAGATGTCTTCTTGT 1256
Db 361 GCACTCCCTGACGAAATCCGAGACAGTCTTCTGAGCTTCTGAGATGTCTTCTTGT 420
QY 1257 GGTCCCTGACATGATCAAGCTGATATCAAGAGATGCTGTGACCTGTCTTCTTCT 1316
Db 421 GGTCCCTGACATGATCAAGCTGATATCAAGAGATGCTGTGACCTGTCTTCTTCT 480
QY 1317 TGAATTTCCGACACCGGCTCAGTGTCTTGAAGACAGAGCCGGCTTTTGTCAAGCCGA 1376
Db 481 TGAATTTCCGACACCGGCTCAGTGTCTTGAAGACAGAGCCGGCTTTTGTCAAGCCGA 540
QY 1377 CCAAGCTGATGAAGTCCGCTTGTGTGCTGAGTGGCTTCTGAAGGGGAGCAATGTAT 1436
Db 541 CCAAGCTGATGAAGTCCGCTTGTGTGCTGAGTGGCTTCTGAAGGGGAGCAATGTAT 600
QY 1437 GTTGAAGAGCCAGGAGAGAGAGTACTGAGCCGAGAGATGATGAAATATCTGGGC 1496
Db 601 GTTGAAGAGCCAGGAGAGAGAGTACTGAGCCGAGAGATGATGAAATATCTGGGC 660
QY 1497 TACCTTGTGGAACCGGGAATCTTAAATGGAACGACCTGTCTGTGGCCAGCTTATTA 1556
Db 661 TACCTTGTGGAACCGGGAATCTTAAATGGAACGACCTGTCTGTGGCCAGCTTATTA 720
QY 1557 TCTGACTGAGAGTACTCTCAGCTGAGCTTGAACATGAGCTTGGACAGAGACTCAAGA 1616
Db 721 TCTGACTGAGAGTACTCTCAGCTGAGCTTGAACATGAGCTTGGACAGAGACTCAAGA 780
QY 1617 ACCGCGGATGATTTTGTGACAGACCAATCCCTGATCTGTGCTCGACATGCT 1676
```

```

DB 781 ACGGGGGGATTTTGGACGACGATCCCGGATCCGCTGCTCGACATGCT 840
QY 1677 CCACAGCTCTCTTCTTCTTAACTTCTCTCTCTCTCCACGCTTTCTTTTCTTTG 1736
DB 841 CCACAGCTCTCTTCTTCTTAACTTCTCTCTCTCTCCACGCTTTCTTTTCTTTG 900
QY 1737 TGCTCCTTGA 1746
DB 901 TGCTCCTTGA 910

RESULT 15
AAD33344
ID AAD33344 standard, cDNA, 921 BP.
XX
AC AAD33344;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human carboxylesterase-like enzyme cDNA.
XX
KM Human; carboxylesterase-like enzyme; organophosphorus intoxication; gene;
KM osteopathic; gene therapy; osteoporosis; antisense therapy; cytosolic;
KM detoxifying agent; Paget's disease; bone implant degradation; cancer;
KM dental implant; enzyme; gene expression; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..921
FT /tag= a
FT /product= "Human carboxylesterase-like enzyme"
FT
PN WO200206454-A2.
XX
PD 24-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-EP007919.
XX
PR 17-JUL-2000; 2000US-0218564P.
XX
PA (PARB ) BAYER AG.
XX
PI Xiao Y;
XX
WI; 2002-195808/25.
DR P-PSDB; AAE20909.
XX
PT Novel human carboxylesterase-like enzyme polypeptide, regulators of which
PT are useful for preventing and treating organophosphorus intoxication,
PT cancer and osteoporosis.
XX
PS Claim 1; Fig 4; 92pp; English.
XX
XX The invention relates to a purified human carboxylesterase-like enzyme
CC polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
CC screening for agents which decrease or modulate the activity of
CC carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
CC useful for treating a carboxylesterase-like enzyme dysfunction related
CC disease, such as organophosphorus intoxication, cancer and osteoporosis.
CC Compounds that increase the ability of human carboxylesterase-like enzyme
CC to bind to organophosphorus compounds are useful as detoxifying agents.
CC Carboxylesterase-like enzyme agonists and antagonists are useful for
CC treating osteoporosis. Paget's disease and degradation of bone implants,
CC particularly dental implants. Carboxylesterase-like enzyme is useful in
CC diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the nucleic acid sequences which encode the enzyme. The
CC coding sequence of carboxylesterase-like enzyme polynucleotide is useful
CC in gene therapy and for generating antisense oligonucleotides or
CC ribozymes which specifically bind to mRNA transcribed from
CC carboxylesterase-like enzyme DNA. These antisense oligonucleotides are

```

```

CC useful for modulating carboxylesterase-like enzyme gene expression. The
CC present sequence is human carboxylesterase-like enzyme cDNA
XX
SQ Sequence 921 BP; 206 A; 238 C; 237 G; 240 T; 0 U; 0 Other;
Query Match 36.3%; Score 634.6; DB 6; Length 921;
Best Local Similarity 86.3%; Pred. No. 1,4e-176;
Matches 743; Conservative 0; Mismatches 49; Indels 69; Gaps 1;
QY 434 CCGTCTGTGTGTGTTCCAGAGAGTGCCTTCAAGACGTGGCTCAGCTCCATCTTGAG 493
DB 56 CAGTTTGTGTGTGTTCCAGAGAGTGCCTTCAAGACGTGGCTCAGCTCCATCTTGAG 115
QY 494 GGTCCGCCCTGCTGCTATGAGACGTGTGTTGCTGCTCAGTACCGGCTAGGAA 553
DB 116 GGTCTCTCTGCTGCTATGAGACGTGTGTTGCTGCTCAGTACCGGCTAGGAA 175
QY 554 TATTTGTTTCTTCAACACATGAGATCAGCATGCTCCGGGAACTGGGCTTCAAGACC 613
DB 176 TATTTGCTTCTTCAACACATGAGATCAGCATGCTCCGGGAACTGGGCTTCAAGACC 235
QY 614 AGGTGCTGCTCTGTCTGTGGTCCAGAGAAATGAGATTTCTTGGTGGGACCCAGCT 673
DB 236 AGGTGCTGCTCTGTCTGTGGTCCAGAGAAATGAGATTTCTTGGTGGGACCCAGCT 295
QY 674 CTGTGACCATCTTTGGCGAGTCCGCGGAGCCATAGTGTCTTATCTATCTGTCTC 733
DB 296 CTGTGACCATCTTTGGTGTAGTCCGTGGAGCCATAGTGTCTTATCTATCTGTCTC 355
QY 734 CCATGGCCAAAGCTTATTTCAAAAGCCATGAGAGTGGGTGGCCATCTATCTCTT 793
DB 356 CCATGGCCAAAGCTTATTTCAAAAGCCATGAGAGTGGGTGGCCATCTATCTCTT 415
QY 794 ACCTGAAGCCCATGATTTATGAAAGATGAGACCTGAGGTGGTGGACATTTCTGTG 853
DB 416 ACCTGAAGCCCATGATTTATGAAAGATGAGACCTGAGGTGGTGGACATTTCTGTG 475
QY 854 GTTAAAGAGCTCAGACTGTAGGCTGTAGGCTGTAGGCTGTAGGCTGTAGGCTGT 913
DB 476 GGTAAAGAGCTCAGACTGTAGGCTGTAGGCTGTAGGCTGTAGGCTGTAGGCTGT 535
QY 914 AGCTGTGACCTCTGACCCAGAAACAAAGCTTTTCACTGAGTGTGTGATGTGCTTCT 973
DB 536 AGCTGTGACCTCTGACCCAGAAACAAAGCTTTTCACTGAGTGTGTGATGTGCTTCT 595
QY 974 TTCTTAATGAGCTCTGATCTTATGTCTCAAGAAACATTTTAAAGCAATTCCTTCATCA 1033
DB 596 TTCTTAATGAGCTCTGATCTTATGTCTCAAGAAACATTTTAAAGCAATTCCTTCATCA 655
QY 1034 TCGAGTCAATTAACACAGAGTGTGCTTCTGCTGCTTGAAGAGGCTCCTGAGATCC 1093
DB 656 TCGAGTCAATTAACACAGAGTGTGCTTCTGCTGCTTGAAGAGGCTCCTGAGATCC 694
QY 1094 TCGAGTCAATTAACACAGAGTGTGCTTCTGCTGCTTGAAGAGGCTCCTGAGATCC 1153
DB 695 -----TGCATCTCTG 706
QY 1154 CTCAGTATTGCACTTGTGCTTAATGATCTTCCATGACAGACCTCCGACTGAA 1213
DB 707 CCAATATTATGCACTTGTGCTGCTTAATGATCTTCCATGACAGACCTCCGACTGAA 766
QY 1214 TCGAGACAGCTTCTGCACTTGTGAGATGTCTTGTGAGTCCCTGACATGATCA 1273
DB 767 TCGAGACAGCTTCTGCACTTGTGAGATGTCTTGTGAGTCCCTGACATGATCA 826
QY 1274 CAGCTCATATCAAGAGATG 1294
DB 827 CAGCTCATATCAAGAGATG 847

```

Search completed: June 14, 2005, 18:10:11
Job time : 944.402 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: June 14, 2005, 23:03:41 ; Search time 127.014 Seconds
(without alignments)
10633.187 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 3170
Sequence: 1 atgcacacaggagactacttc.....ttctcttctgtgcctctga 1746

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xln
-Q=/cg2_1/USPTO.spool/6664091/rnat_14062005_133458_9325/app.query.fasta.1.4238
-DB=A.GeneSeq_16Dec04 -OPMT=faetan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000
-USER=6664091 @CGN_1_1_299 @rnat_14062005_133458_9325 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_GeneSeq_16Dec04: *
1: geneeqp1980s: *
2: geneeqp1990s: *
3: geneeqp2000s: *
4: geneeqp2001s: *
5: geneeqp2002s: *
6: geneeqp2003as: *
7: geneeqp2003bs: *
8: geneeqp2004s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3079	97.1	581	5	ABB79537 Human car
2	3079	97.1	581	8	ADQ89094 Human uro
3	2915	92.0	575	5	ABP61004 Human pro
4	2900	91.5	642	5	AAE25025 Human dia
5	2900	91.5	642	8	ABM84114 Human dr
6	2607.5	82.3	581	5	ABP61005 Human hum
7	2606	82.2	525	6	ADA54807 Human pro
8	2597.5	81.9	581	6	ABU54639 Human NOV
9	2588.5	81.7	618	5	ADR19663 Human dru
10	2454	77.4	469	7	ADB64065 Human pro

11	1985	62.6	542	8	ADP50145	Adf50145 Cat cauxi
12	1985	62.6	542	8	ADP50147	Adf50147 Cat cauxi
13	1780	56.2	356	5	ABP61006	ABP61006 Novel hum
14	1262.5	39.8	561	5	AAE20911	AAE20911 Rat carbo
15	1230	38.8	549	3	AAE58981	AAE58981 Breast an
16	1230	38.8	550	8	ADG98219	ADG98219 Human int
17	1230	38.8	559	4	AAE31700	AAE31700 Protein e
18	1230	38.8	559	4	ABU53223	ABU53223 Human met
19	1230	38.8	559	6	ABR82851	ABR82851 CES2 rela
20	1230	38.8	559	6	AAE32960	AAE32960 Human car
21	1230	38.8	559	7	ADD46521	ADD46521 Human pro
22	1221	38.5	306	5	AAE20909	AAE20909 Human car
23	1213.5	38.3	561	7	ADD46519	ADD46519 Rat Prote
24	1191	37.6	607	4	ABU53222	ABU53222 Human met
25	1186.5	37.4	554	8	AD179890	AD179890 Mouse liv
26	1181	37.3	583	4	ABG10273	ABG10273 Novel hum
27	1179.5	37.2	571	3	AAV71107	AAV71107 Human Hyd
28	1179.5	37.2	571	4	AAU12442	AAU12442 Human PRO
29	1179.5	37.2	571	4	AAE04101	AAE04101 Human gen
30	1179.5	37.2	571	5	AAU83696	AAU83696 Human PRO
31	1179.5	37.2	571	5	ABR84949	ABR84949 Human PRO
32	1179.5	37.2	571	5	ABG64341	ABG64341 Human alb
33	1179.5	37.2	571	6	ABO17886	ABO17886 Novel hum
34	1179.5	37.2	571	6	ABU65108	ABU65108 Human PRO
35	1179.5	37.2	571	6	ABU80843	ABU80843 Human PRO
36	1179.5	37.2	571	6	ABO33809	ABO33809 Novel hum
37	1179.5	37.2	571	6	ABU81140	ABU81140 Human PRO
38	1179.5	37.2	571	6	ABO19424	ABO19424 Human sec
39	1179.5	37.2	571	6	ABU66840	ABU66840 Human PRO
40	1179.5	37.2	571	6	ABU59921	ABU59921 Novel sec
41	1179.5	37.2	571	6	ABU69085	ABU69085 Human PRO
42	1179.5	37.2	571	6	ABO25111	ABO25111 Human sec
43	1179.5	37.2	571	6	ABU82152	ABU82152 Novel hum
44	1179.5	37.2	571	6	ABU67116	ABU67116 Human sec
45	1179.5	37.2	571	6	ABU81549	ABU81549 Human sec

ALIGNMENTS

RESULT 1	
ABB79537	standard; protein; 581 AA.
ID	ABB79537
AC	ABB79537;
XX	
XX	23-SEP-2002 (first entry)
DT	
XX	
DE	Human carboxylesterase family member 53010.
XX	
KW	Carboxylesterase; enzyme; human; analgesic; nootropic; antiinflammatory;
KW	diagnosis; therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	/label= Signal_peptide
FT	Protein
FT	/label= Mature_protein
FT	Domain
FT	/label= Mature_protein
FT	Region
FT	/note= "carboxylesterase domain"
FT	Active-site
FT	/note= "predicted carboxylesterase type-B signature"
FT	/note= "predicted carboxylesterase type-B serine active site"
XX	
XX	MO200250256-A2.
XX	
XX	27-JUN-2002.
XX	
XX	18-DEC-2001; 2001WO-US049075.
XX	

QY 1501 TTGCTGGAACGGGAATCTTAATGGGAACGACCTGCTCTGTGGCCAGCTTAATCTG 1560
DB 501 PhalaiahglnrGlyAsnProasnGlyAsnAspLeuSerLeuTrpProlAlaTrpAsnLeu 520
QY 1561 ACTGAGCAGTACCTTCACAGTGAAGTGAAGCTCGGACAGAGACTCAAGACCG 1620
DB 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlnGlnArgLeuGlnPro 540
QY 1621 CGGCTGATTTTGTGACCAAGACCATCCCTGATCCCTGCTGCTCCGACATGCTCAC 1680
DB 541 ArgValAspPheTrpThrSerThrLeuProLeuLeuSerAlaSerAspMetLeuHis 560
QY 1681 AGTCTCTCTTCTCTTCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 561 SerProLeuSerSerLeuTrpPheLeuSerLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
DB 581 Pro 581
RESULT 2
ADQ89094
ID ADQ89094 standard; protein; 581 AA.
XX
AC ADQ89094;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human urological disorder related protein 53010 SEQ:46.
XX
KM urological disorder; uropathic; cytostatic; urinary incontinence;
KM benign prostatic hyperplasia; human.
XX
OS Homo sapiens.
XX
PN WO2004065576-A2.
XX
PD 05-AUG-2004.
XX
PE 14-JAN-2004; 2004WO-US000750.
XX
PR 15-JAN-2003; 2003US-0440318P.
PR 04-FEB-2003; 2003US-0444783P.
PR 27-MAR-2003; 2003US-0457901P.
PR 08-MAY-2003; 2003US-0468775P.
PR 19-MAY-2003; 2003US-0471614P.
PR 16-JUN-2003; 2003US-0478742P.
PR 18-JUL-2003; 2003US-0488523P.
PR 30-JUL-2003; 2003US-0491156P.
PR 02-SEP-2003; 2003US-0495949P.
PR 26-SEP-2003; 2003US-0506332P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Karicheti V, Silos-Santiago I, Eliasof SD;
XX
XX WPI; 2004-562167/54.
DR N-PSDB; ADQ89093.
XX
PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
PT 211 or for identifying a compound capable of treating a urological
PT disorder or identifying and treating a subject having a urological
PT disorder.
XX
XX
PS Claim 1; SEQ ID NO 46; 542pp; English.
CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a
CC urological disorder, identifying a subject having a urological disorder,
CC or treating a subject having a urological disorder. Also described: (1) a
CC method for identifying a compound capable of treating a urological
CC disorder; (2) a method for identifying a subject having a urological

CC disorder; and (3) a method for treating a subject having a urological
CC disorder. The compound has uropathic and cytostatic activities. The
CC polypeptides related to urological disorders are useful for identifying a
CC compound capable of treating a urological disorder, identifying a subject
CC having a urological disorder, or treating a subject having a urological
CC disorder. Disorders include urinary incontinence and benign prostatic
CC hyperplasia. The present sequence represents a human urological disorder
CC related protein, which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 581 AA;
Alignment Scores:
Pred. No.: 3,216-287 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 97.13% Indels: 0
DB: Gaps: 0
US-10-023-515-3 (1-1746) x ADQ89094 (1-581)
QY 1 ATGCCACAGGAGACTTACTTCACTGCTTCAACATGTCCTTTTCTGATTCACGCC 60
DB 1 MetProGlnGlyLeuTrpSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 20
QY 61 CTGTTGGGACACACAGTGGGGAACAACTGGGCTTCTGCTGAAGGCGCACAGAGAAC 120
DB 21 LeuLeuGlnHisArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgAsn 40
QY 121 ACCAGGCTGGATGATTCAGGAGCAGCAAGTCACTGCTGAGGAAGCCCTGCTGTG 180
DB 41 ThrArgLeuGlnTrpIleGlnGlyLeuGlnValThrValLeuGlnSerProValProVal 60
QY 181 AACGTGTTCTCTGAGAGTCCCTTTGCTGCTCCCGCTGGAGATCCCTGCAATTAAGAAC 240
DB 61 AsnValPheLeuGlnValProPheAlaAlaProProLeuGlnSerLeuArgPheThrAsn 80
QY 241 CCGCAGCTGATGCGCTTGGGATTAATCTTGGGAAGCAACCTCTCAATTTGGC 300
DB 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnAlaTrpSerTyrProAsnLeuGly 100
QY 301 CTCGAGAACTCAGAGTGGCTGCTCTTGAATCAACATGCTCAAGAGTCATTACCGGAA 360
DB 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisSheLeuValHisTyrProGly 120
QY 361 TTGGAAGTGCAGAAAGCTGCTTACCTGAACATTAAGCGCTGCCACGCGGATACA 420
DB 121 PheGlyValSerGlnuAspCysLeuTrpLeuAsnIleTyrAlaProAlaHisAlaAspThr 140
QY 421 GGTTCGAAGCTCCCGCTTGGTGTGGTTCCTCCAGAGAGTGGCTTCAGACTGGCTCAGCC 480
DB 141 GlySerIysLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySerIle 160
QY 481 TCCATCTTGAATGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 161 SerIlePheAspGlySerAlaLeuAlaAlaTyrGlnuAspValLeuValValAlaGln 180
QY 541 TACCGGCTAGGAATATTTGTTCTTCAACCATGAGATCAGATGCTCCGCGGAATCGG 600
DB 181 TyrArgLeuGlnIlePheGlyPhePheThrTrpAspGlnHisAlaProGlyAsnTrp 200
QY 601 GCTTCAAGACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 AlaPheIysAspGlnValAlaAlaLeuSerTrpValGlnuAsnIleGlnPhePheIly 220
QY 661 GGGAGCCCGAGCTGAGACCATCTTTGGCAGAGCCGCGGAGCCATTAAGTTTCTAGT 720
DB 221 GlyAspProSerSerValThrIlePheGlyGlnuSerAlaGlyAlaIleSerValSerSer 240
QY 721 CTTAATCTGCTTCCATGCGCAAGGCTTAATTCACAAAGCCATCATGAGAGTGGGCTG 780
DB 241 LeuIleuSerProMetAlaIleGlyLeuPheHisIlyAsaIleMetGlnuSerGlyVal 260

```

QY 781 GCATCATCCCTTACCTGGAGGCCATGATTATGAGAGAGTGGAGCACTGGAGGTGTT 840
DB 261 AATleilelProGlyLeuGluAlaHisAspTyrgLyluSerGluAspLeuGlnAlaVal 280
QY 841 GCACATTTGTGTGTATACATAGCGTGAAGCTTGAGAGCCCTGCTGAGTGGCTGAGACA 900
DB 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuAlaGlySerLeuArgThr 300
QY 901 AAACCTCCAGAGAGCTGTGACCCCTGACCCAGCAAAACAAAGCTTTTCACTGAGTGGTT 960
DB 301 LysProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
QY 961 GATGAGCTTTTCTTCTATAGAGCTCTAGATCTTATGTCCTCAGAAAGCATTTTAAAGA 1020
DB 321 AspGlyAlaPhePheProAsnIuProLeuAspLeuLeuSerGlnAlaPheLysAla 340
QY 1021 ATTCCTTCATATATGAGAGTCAATTAACAAGAGTGTGCTTCTGCTGCTATGAGAGAG 1080
DB 341 IleProSerIleIleGlyValAsnAsnHisGluCysGlyPheLeuLeuProMetLysGlu 360
QY 1081 GCTCCAGATGCTCTAGTGGCTCCAAAGAGCTTGGCCCTGCTGATATCAAAACATC 1140
DB 361 AlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLysLysIleGlnAsnIle 380
QY 1141 CTGACATCCCGCCTCAGTATTTTGACACCTTGCTGCTAAATGATCTTCATGACAAAGCAC 1200
DB 381 LeuHisIleProProGlnTyrluThrLysLeuValAlaAsnGlnTyrluPheHisAspLysHis 400
QY 1201 TCCCTGATGTAATCCGAGACAGAGTCTTGAGCTTGTGAGATGTTCTTTGTGCTC 1260
DB 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValAla 420
QY 1261 CCTGACATGATCAGAGCTGATATCAAGAGAGTGTGCTGAGACCTGTACTTCTATGAG 1320
DB 421 ProAlaLeuIleThrAlaArgTyrluHisArgAspAlaGlyAlaProValTyrluPheSerGlu 440
QY 1321 TTTCGGGACCGGCTCAGTGTCTTTGAGAGACAGAAAGCCGCTTTGTCAAAAGCCGACAC 1380
DB 441 PheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHis 460
QY 1381 GCTGATGAAAGTCCCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 461 AlaAspGluValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
QY 1441 GAAGAGACCAAGAGAGAGAGAGTACTGAGCCGAGAAAGATGATGAATACTGGGCTACC 1500
DB 481 GluGlyAlaThrGlnGluGlnLysLeuLeuSerArgLysMetMetLysTyrluPheAlaThr 500
QY 1501 TTTCCTGAAACCGGGAATCTTAATGGAAAGCACTGTCTCTGTGGCCAGCTTATATCTG 1560
DB 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTyrluProAlaTyrluAsnLeu 520
QY 1561 ACTGAGAGATACCTCCAGCTGAGATTGAACATGAGCTCCGAGAGAGACATCAAGAAAGCCG 1620
DB 521 ThrGlnGlnTyrluLeuGlnLeuAspLeuAsnMetSerLeuGlnAlaGlyLeuLysGluPro 540
QY 1621 CGAGTGAGATTTTGGACACAGACCATCCCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 541 ArgValAspPheThrPheThrSerThrIleProLeuIleLeuSerHisAspAspMetLeuHis 560
QY 1681 AGTCCCTTTTCTTCTTAACCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
DB 581 Pro 581

```

RESULT 3
ABP61004
ID ABP61004 standard; protein; 575 AA.
XX

```

AC ABP61004;
XX 10-SEP-2002 (First entry)
DT
XX
DE Novel human protein. SEQ ID 91.
XX
KW Human; cytosol; vulnary; antiarteriosclerotic; antiparkinsonian;
KW neurotrophic; neuroprotective; immunosuppressive; haemostatic;
KW antinflammatory; cardiant; antitumor; vitruide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorder; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological diseases; metabolic diseases; sperm dysfunction; thyroid
KW disorders e.g. hypothyroidism; brain damages; colitis; cone photo-
KW transduction deficiency; neurological diseases; stroke; anglogenesis,
XX
OS Homo sapiens.
XX
XX WO000250105-A1.
XX
PD 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.
XX
PR 20-DEC-2000; 2000US-0257048P.
XX
PR 09-JAN-2001; 2001US-0260482P.
XX
PR 30-JAN-2001; 2001US-0264922P.
XX
PR 06-FEB-2001; 2001US-0266797P.
XX
PR 19-MAR-2001; 2001US-0276988P.
XX
PR 04-APR-2001; 2001US-0281535P.
XX
PR 08-MAY-2001; 2001US-0289622P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkealand M, Cogswell JP, Kabnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX WPI; 2002-508784/54.
XX N-PSDB; ABQ86169.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 1(a); Page 312-313; 335pp; English.
XX
XX
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytotoxic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotrophic,
CC neuroprotective, immunosuppressive, haemostatic, antinflammatory,
CC cardiant, antitumor, vitruide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC cardiovascular disorders, pancreatitis, nervous system related disorders and
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,

```

CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention

XX Sequence 575 AA:

Alignment Scores:

Pred. No.:	2,096-271	Length:	575
Score:	2915.00	Matches:	553
Percent Similarity:	97.54%	Conservative:	3
Best Local Similarity:	97.02%	Mismatches:	6
Query Match:	91.96%	Indels:	8
DB:	5	Gaps:	1

US-10-023-515-3 (1-1746) x ABP61004 (1-575)

```

QY 34 TGTGCTTTTTCGATTCCTCCAGCCCTGTTGGACACAGAGTGGGAAACCTGG 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14 TtrpAlaIleTtrpValIleuAlaIlePro-----ThrlYsGly 25

QY 94 CTTTCTGCTGAAGGCGCCACAGAGAACACAGCTGGATGAGTTCAGGCGAACAGCTC 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26 ProSerAlaGluGlyProGlnArgAanThraGleuGlyTrrpIleGlnGlyYsGlnVal 45

QY 154 ACTGTGCTGGGAAGCGCTGCTGTAACGTTCTCCGAGTCCCTTGGCTGCTGCC 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46 ThrValIleuGlySerProValProValAanValPheIleuGlyValProPheIleAlaIlePro 65

QY 214 CCGCTGGGATCCCTGCGATTAACGAACCCGACGCTGATGCGCCCTGGGATTACTTGCA 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 ProIleuGlySerIleuArgPheThraSerProGlnProAlaSerProTrrpAspIleuArg 85

QY 274 GAAGCCACTCTCTACCTTAATTTTGCTCCAGAACTCAGAGTGGCTCTTATGATCAA 333
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 GluAlaIleThraSerTrrpProAsnIleuGlyIleuGlnAanSerGluTrrpIleuIleuAspGln 105

QY 334 CACATGCTCAAGTGTGCAATTACCCGAATTCGAGTGTCAAGAACTGCTCTACCTGAC 393
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 106 HisIleuIleuValIleValIleValIleValIleValIleValIleValIleValIleVal 125

QY 394 ATCTATGCGCTGCGCCACGCGCATCAGAGCTCAAGCTCCCGTCTTGATGTTGCCA 453
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 IleThraIleProAlaHisIleAlaSerThrIleSerIleuProValIleuValIlePhePro 145

QY 454 GAGAGTGCCTTCAAGACTGCTCAGCTCACTTGTATGGTCCGCTGGCTGCTCTAT 513
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 GlyIleAlaPheIleThrlGlySerAlaSerIlePheAspGlySerAlaIleuAlaIleAlaTyr 165

QY 514 GAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 573
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 GluAspValIleuValIleValIleValIleValIleValIleValIleValIleValIleVal 185

QY 574 TGGATGACGATGCTCGGGGAACTGGGCTTCAAGGACAGGATGCTGTGCTGCTG 633
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 TrpAspGlnHisIleAlaProGlyAanTrrpAlaPheIleuAspGlnValAlaIleuSerTrrp 205

QY 634 GTCCAGAGAACATCGAGTCTTTCGATGAGGACCCCACTGTGTGACCATCTTGGCGAG 693
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 ValGlnIleuValIleGluPhePheGlyIleYAspProSerSerValThrllePheGlyGlu 225

QY 694 TCCGCGGAGGACATTAAGTGTTCATGCTTATCTGTCTCCCATGCGCAAGCTTATTC 753
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 SerIleuGlyAlaIleSerValSerIleuIleuSerProMetAlaIleGlyIleuPhe 245

QY 754 CACAAAGCCATCATGAGAGTGGGATGAGCATCATCCCTTACCTGAGGACCATATTTAT 813
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 HisIleuAlaIleIleuGlnSerIleValAlaIleIleProTrrpIleuGlnAlaHisIlePrr 265

QY 814 GAGAGAGTGAAGACTTCAAGTGTGTGACATTTCTGTGTGTACATGCGTCAACTCT 873
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 GluIleuSerGlnAspIleuGlnValIleAlaHisPheCysGlyAanAsnAlaSerAspSer 285
  
```

```

QY 874 GAGGCCCTGTGAGGTGCTGAGGACAAACCTTCAAGAGACTGTGACCCCTGAGCCAG 933
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 GluAlaIleuIleuArgCysIleuArgThrlYsProSerIleuGlnIleuThrlleuSerGln 305

QY 934 AAAACAAAGCTTTTCACTCGAGTGTGATGTGCTTTCTTCTTATGACCTTACAT 993
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 LysThrlYsSerPheThraArgValIleAspGlyAlaPhePheProAsnGlnProIleuAsp 325

QY 994 CTATTGTCTCGAAGCAATTTAAAGCAATTCCTTCCATCATCGGAGTCAATACACAGAG 1053
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 LeuIleuSerGlnIleuAlaPheIleAlaIleProSerIleIleGlyValAanAsnHisGln 345

QY 1054 TGTGCTTCCGCTGAGCTTATGAAGAGGCTCTGAGATCTCACTGAGCTCCCAAGTCC 1113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 CysGlyPheIleuIleuProMetIleGlnAlaProGluIleuSerGlySerAsnIleuSer 365

QY 1114 CTTCCTCCATCTGATACAAACATCTCGACATCCCGCTCAGATATTGCACTGTGTG 1173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 LeuAlaIleuHisIleuIleGlnAanIleuHisIleProProGlnTrrpIleuHisIleuVal 385

QY 1174 GCTAATGAATCTTCCATGACAAAGCACTCCCTGACTGAATCCGAGACAGCTTCTGGAC 1233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 AlaAsnGluTrrpPheHisAspIleHisSerIleuThrlGluIleArgAspSerIleuAsp 405

QY 1234 TTGCTTGGAGATGATGTTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1293
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 LeuIleuGlyAspValPhePheValIleProAlaIleuIleThrlAlaArgTrrpHisIleArg 425

QY 1294 GCTGTGCACTGTCTACTTCTATGATGTTTGGGACCGGCTCAGATGCTTGAAGACAG 1353
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 AlaGlyAlaProValTrrpPheTrrpGluPheArgHisArgProGlnCysPheGluAspThr 445

QY 1354 AAGCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTGCTGCTGCC 1413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 LysProAlaPheValIleValAspHisIleAlaAspGlnValArgPheValPheGlyGlyAla 465

QY 1414 TTCCGAAAGGGGACATGTTATGTTTGAAGAGACACGAGAGAGAGAAATTTCTGAGC 1473
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 PheIleuIleuGlyAspIleValIleMetPheGluIleAlaThrlGluGlnIleuIleuSer 485

QY 1474 CGGAAGATGATGAATATCTGGGCTACCTTGTCTGAAACCGGGAATCTTATGGAGAC 1533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 ArgIleuMetIleuTrrpTrrpAlaThraPheAlaArgThrlGlyAanProAsnGlyAanAsp 505

QY 1534 CTGTCTGTGCGGACCTTAATATCTGACTGAGACAGTACTCTGAGCTGGAACATG 1593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 LeuSerIleuTrrpProAlaTrrpAsnIleuThrlGluIleuGlnIleuAspIleuAsnMet 525

QY 1594 AGCCTCGGACAGAGACTCAAGAAACCGGAGTGGATTTTGGACACAGACATCCCTG 1653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 526 SerIleuGlyIleuValIleGluIleuValIleProArgValAspPheTrrpHisIleProIleu 545

QY 1654 ATCTGTCTGCTCGACATGCTCCACAGTCTCTTCTTCTTAACCTTCTCTCTC 1713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 IleIleuSerAlaSerAspMetIleuHisSerProIleuSerSerIleuThraPheIleuSerIleu 565

QY 1714 CTCAGCTTTCTTTTCTTTTCTTTGCTCT 1743
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 LeuGlnProPhePhePheCysAlaIlePro 575

RESULT 4
AAE25025
ID AAE25025 standard; protein; 642 AA.
AC AAE25025;
XX
XX
DT 30-OCT-2002 (first entry)
XX
XX Human drug metabolizing enzyme (DME-10).
XX
XX Human; drug metabolizing enzyme; autoimmune; inflammatory disorder;
KW acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;
  
```

KW proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;
KW asthma; neurological disorder; Alzheimer's disease; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; anemia; adenoma;
KW drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;
KW renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;
KW anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;
KW goitre; gastrointestinal disorder; gene therapy; viroicide; anticoagulant;
KW anticonvulsant; noctropic; enzyme; DME-10.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..50
FT /label= Signal_peptide
FT Domain 32..56
FT /note= "Transmembrane domain"
FT Protein 51..642
FT /note= "Mature human DME-10"
FT Domain 113..135
FT /note= "Transmembrane domain"
FT Domain 204..220
FT /note= "Transmembrane domain"
FT Domain 234..250
FT /note= "Transmembrane domain"
FT Domain 287..314
FT /note= "Transmembrane domain"
FT Domain 463..491
FT /note= "Transmembrane domain"
XX
PN MO200246426-A2.
XX
PD 13-JUN-2002.
XX
PF 04-DEC-2001; 2001MO-US047429.
XX
PR 08-DEC-2000; 2000US-0254308P.
PR 15-DEC-2000; 2000US-0256189P.
PR 21-DEC-2000; 2000US-0257713P.
PR 19-JAN-2001; 2001US-0262706P.
PR 02-FEB-2001; 2001US-0266020P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,
PI Lee EA, Ding L, Hafalia AJA, Tang YT, Yue H, Tribouley CM, Lu DM,
PI Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y,
PI Ieon CH;
XX
DR WPI: 2002-519668/55.
XX N-PSDB: AAD40574.
XX
PT Novel human drug metabolizing polypeptide, useful in diagnosis,
PT prevention or treatment of autoimmune/inflammatory, cell proliferative,
PT neurological, developmental, endocrine, metabolic and gastrointestinal
PT disorders.
XX
PS Claim 65; Page 155-156; 169pp; English.
XX
CC The invention relates to an isolated human drug metabolizing enzyme (DME)
CC and its nucleotide. DME is useful for diagnosing, treating or preventing
CC disorders associated with aberrant expression of DME, where the disorders
CC are selected from autoimmune/inflammatory disorder such as acquired
CC immunodeficiency syndrome (AIDS), asthma, arteriosclerosis, psoriasis,
CC liveritis; a cell proliferative disorder such as arteriosclerosis,
CC cirrhosis, hepatitis, and cancer; a neurological disorder such as
CC Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;
CC a developmental disorder such as renal tubular acidosis, epilepsy,
CC anaemia; an endocrine disorder such as adenoma, thrombosis and infections
CC ; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic
CC disorder such as cystic fibrosis, diabetes and goitre; a gastrointestinal
CC disorder such as anorexia, peptic ulcer; and liver disorders. DME is
CC useful in a number of drug screening techniques and to analyse the
CC proteome of a tissue or cell type. The invention is useful for creating

CC knock-in humanised animals or transgenic animals to model human diseases,
CC in somatic or germline gene therapy, to generate a transcript image of a
CC tissue or cell type, for detecting differences in the chromosomal
CC location due to translocation, inversion, etc. among normal carrier or
CC affected individuals, and as hybridisation probes for mapping naturally
CC occurring genomic sequences. The present sequence is human DME-10
XX
SQ Sequence 642 AA;
SQ
Alignment Scores:
Pred. No.: 6,21e-270 Length: 642
Score: 2900.00 Matches: 548
Percent Similarity: 99.82% Conservative: 2
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 91.48% Indels: 0
DB: 5 Gaps: 0
US-10-023-515-3 (1-1746) x AAE25025 (1-642)
QY 91 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGCGTGGATGATTCAAGGCAAGCAA 150
DB 92 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLeuGln 111
QY 151 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAAGTGTCTCGAGATCCCTTGTCTGCTGCT 210
DB 112 ValThrValLeuGluGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 131
QY 211 CCCCCTGGGAATCCCTGCAGATTACGAACCCGACGCTGCATCGCCCTGGATTAACCTTG 270
DB 132 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpPheAsnLeu 151
QY 271 CGAAGAGCCACCTCCCTACCTTAATTTGTGCTCTCAAGACTCAGAGTGGCTCTTAGAT 330
DB 152 ArgGlnIleThrSerLeuProAsnLeuGlyLeuGlnAsnSerGlyLeuProLeuLeuAsp 171
QY 331 CAACACATGCTCAAGGTGATTAACCCGAATTCGAGTGCAGAAAGCTGCCCTTACCTTG 390
DB 172 GlnHisMetLeuLeuSerValHisTrpProLysPheGlyValSerGlnAspCysLeuTrpLeu 191
QY 391 AACATCTATGCGCTCGCCCAACGCCGATACAGGCTCCAGCTCCCGCTCTGTGGTTC 450
DB 192 AsnIleTrpAlaProAlaHisAlaPheThrGlySerLeuProValLeuValTrpPhe 211
QY 451 CCAGAGAGTGCCTTCAAGCTGCTCAGCTCCATCTTGAATGGGTCCGCTGCTGCTGCC 510
DB 212 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 231
QY 511 TATGAGACGTGCTGTGTGTGTGCTGCTCAGTACCGGCTAGAAATATTGGTTCTTCACC 570
DB 232 TyrGluAspValLeuValValValValGlnTrpArgLeuGlyIlePheGlyPhePheThr 251
QY 571 ACATGGAGTACACATGCTCCGGGAACTGGGCTTCAAGAACACAGGTGGCTGCTGTGCC 630
DB 252 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaIleLeuSer 271
QY 631 TGGGTCGAAGAAGACATCGATGTTCTTCGGTGGGAGCCCGAGCTCGTACCATCTTGGC 690
DB 272 TrpValGlnIlyAsnIleGluPhePheGlyValAspProSerSerValThrIlePheGly 291
QY 691 GAGTCCGCGGAGGACATTAAGTGTTCATGCTTATACGTCTCCCATGGCCAAAGCTTAA 750
DB 292 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPrometAlaLysGlyLeu 311
QY 751 TTCGCAAAAGCATCATGAGAGTGGGTGGCCATCATCCCTTACTCTGAGGCCCATAT 810
DB 312 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 331
QY 811 TATGAGAAAGTGAAGACCTCAAGGTGTGTGACATTTCTGTGGTAAATGCGTGAGAC 870
DB 332 TyrGluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnHisSerAsp 351
QY 871 TCTGAGGCCCTGCTGAGGTGCTTGAAGACAACCCCTTCAAGAGCTGTGACCTTCAGC 930

```

Db      352 SerGluAlaLeuLeuArgCysLeuArgThrIysProSerIysGluLeuLeuThrLeuSer 371
Qy      931 CAGAAAACAAGCTCTTTCACGAGTGAGTGAGTGGTCTTCTTCTTAAGAGCTCTTA 990
Db      372 GlnIysThrIysSerPheThrArgValAlaIspGlyAlaPhePheProAsnGluProLeu 391
Qy      991 GATCTATTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1050
Db      392 AspleuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAlaAsnHis 411
Qy      1051 GAGTGTGGCTTCTGCTGCTCCATGAGAGAGGCTCCTGAGATCTCAGTGGCTCAACAG 1110
Db      412 GluGlyGlyPheLeuLeuPheProMetIysGlyAlaProGluValLeuSerGlySerAsnIys 431
Qy      1111 TCCCTTGCCCTCCATCGATCAAAACATCCTGCAATCCGCGCCCTCAATTTGACCTT 1170
Db      432 SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIlyrLeuHisLeu 451
Qy      1171 GTGGCTAATGAATCTTCCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTTCTG 1230
Db      452 ValAlaAsnGluIyrPheHisAspIysHisSerLeuThrGluIleArgAspSerLeuLeu 471
Qy      1231 GACTTGTGGAGAGTGTGTCTTGTGGTCCCTGCACTGATCAAGCTCGATATCAACA 1290
Db      472 AspleuLeuGlyAspValPhePheValAlaProAlaLeuIleThrAlaArgIyHisArg 491
Qy      1291 GATGCTGGTGACCTGTCTACTTCTATGAGTTTGGGACCCGCGCTCAGTGGTTGAAC 1350
Db      492 AspAlaGlyAlaProValIyrPheIyIyGluPheArgHisArgProGlnIyPheGluAsp 511
Qy      1351 ACAGAGCCGGCTTTTGTCAAGCCGACGACGCTGATGAAGTCCGCTTGTGTGGTGGT 1410
Db      512 ThrIysProAlaPheValIysAlaAspHisAlaAspGluValArgPheValPheGlyIy 531
Qy      1411 GCCTTCTGAAGGGGACATTGTATGTTCGAAGAGCCACGAGAGAGAGAACTTACTG 1470
Db      532 AlaPheLeuIysGlyAspIleValMetPheGluGlyAlaThrGluGluIyLeuLeu 551
Qy      1471 AGCGGGAAGATGAGAAATACTGGGCTACCTTGTGTCACACCGGGAATCCCTAATGGGAC 1530
Db      552 SerIyIysMetIyIyIyIyIyAlaThrPheAlaArgIyIyIyAsnProAsnIyAsn 571
Qy      1531 GACCTGTCTGTGGGACGCTTATATCTGACTGAGTACGCTCCAGCTGGAATTGTAAC 1590
Db      572 AspleuLeuLeuIyrProAlaIyIyAsnLeuThrGluIyIyLeuGlnIleuAsn 591
Qy      1591 ATGAGCCTCGGACAGAGACTCAAAAGAACCGGCGTGAATTTTGACAGCAATCCCC 1650
Db      592 MetSerLeuGlnIyGlnArgLeuIyGluIyrProArgValGluPheIyrThrSerThrIlePro 611
Qy      1651 CTGATCTGTGTGGCTCCGACAGATGCTCAAGTCTCTTCTTCTTAATTTCTCTCT 1710
Db      612 LeuIleuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 631
Qy      1711 CTCCCTCAGAGCTTTCTTTCTTTTGTGCTCCT 1743
Db      632 LeuLeuGlnIyrProPhePhePhePheCysAlaPro 642

```

```

RESULT 5
ABM84114
ID ABM84114 standard; protein; 642 AA.
XX
AC ABM84114;
XX
XX 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic protein SEQ ID NO:4363.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dthp.
XX
XX Homo sapiens.
XX
XX WO2004023973-A2.

```

```

XX      25-MAR-2004.
PD      12-SEP-2003; 2003WO-US028227.
XX
PF      12-SEP-2002; 2002US-0410259P.
XX
PR      12-SEP-2002; 2002US-0410260P.
XX
XX      (INCY-) INCYTE CORP.
XX
PI      Schmidt JP, Wright RJ, Brune CM, Marianovic MM, Shen F;
PI      Harthorn TA, Suchorolski MT, Altus CM, Plets SO, Elder LV;
PI      Mooney EM, Deleage AM, Panesar IS, Barville SC, Reddy TP;
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI      Peralta CH, Anderson SB, Rioux P, Shen RJ, Wu MC, Stuve LL;
PI      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI      Patury S, Shi X, Suarez CJ;
PI
PI      MPI; 2004-329368/30.
DR      N-PsDB; ACN42766.
XX
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX      in diagnosing a condition, disease or disorder associated with human
XX      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX      in gene mapping.
XX
PS      Claim 27; Page; 190pp; English.
XX
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
XX      selected from one of the 2722 sequences defined in the specification. A
XX      polynucleotide of the invention may have a use in gene therapy. The human
XX      diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
XX      used to diagnose a particular condition, disease or disorder associated
XX      with human molecules, e.g. cell proliferative disorders,
XX      autoimmunity/inflammatory disorders, developmental disorders, endocrine
XX      disorders, neurological disorders, gastrointestinal disorders, or
XX      infections caused by virus, bacteria, fungi or parasite. The dthp
XX      molecules may also be used in genetic mapping, in identifying individuals
XX      from minute biological samples, in detecting single nucleotide
XX      polymorphisms, as molecular weight markers, and for somatic or germline
XX      gene therapy. The present sequence represents a dthp protein of the
XX      CC invention. Note: The sequence data for this patent is not represented in
XX      the printed specification, but was obtained in electronic format directly
XX      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ      Sequence 642 AA;
XX
XX      Alignment Scores:
XX      Pred. No.: 6,21e-270 Length: 642
XX      Score: 2900.00 Matches: 549
XX      Percent Similarity: 99.82% Conservative: 1
XX      Best Local Similarity: 99.64% Mismatches: 1
XX      Query Match: 91.48% Indels: 0
XX      DB: 8 Gaps: 0
XX
US-10-023-515-3 (1-1746) x ABM84114 (1-642)
Qy      91 GGCGCTTCTGTGAAGGGCCACAGAGAAACAGGCTGGATGATTCAAGGCAAGCA 150
Db      92 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyIyrPheIleGlnIyIyAsn 111
Qy      151 GTCACTGTGTGGAGAGCCCTGTGCTGTGAAGCTGTTCCTCGAGAGTCCCTTGTGCTCT 210
Db      112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIle 131
Qy      211 CCCCCTGGGATCCCTCGCATTTACGAACCCGAGAGCTGATGAGCCCTGGGATTAATCTTG 270
Db      132 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProIyrAspAsnLeu 151
Qy      271 CGAGAGCCACCTCTCAATCTTATTTGTGCTCGAGAACTCAGAGTGGCTCTTATGAT 330
Db      152 ArgGluAlaThrSerIyIyrProAsnLeuIyLeuGlnAsnSerGluIyrPheLeuLeuAsp 171

```


QY 331 CAACATGCTCAAGGTCATTACCCGAAATTGAGTGTGAGAGACTGCTTACCTG 390
DB 172 GlnHMeLeuLeuYsAlaHnStYrProLysPheGlyValSerCyluSprCysLeuYrLeu 191
QY 391 AACATGTTATGCGCTGCGCCCAAGGTCAGAGCTCCCGCTCTTGTGTGTGTC 450
DB 192 AamIleYrAlaProAlaHnAlaSerThrGlySerLeuLeuProValLeuValTPrphe 211
QY 451 CAGAGGAGTGTGCTTCAAGACTGCTCAGCTCCATTTTGAATGGGTCCGCTGGCTGCC 510
DB 212 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 231
QY 511 TATGAGACGTGCTGTTGTGCTGCTCAGTACCGGCTAGAAATTTTGTCTTCCACC 570
DB 232 TyrGluAspValLeuValAlaValGlnTyrAlaGlyLeuGlyIlePheGlyPhe***Ser 251
QY 571 ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGCTGTCC 630
DB 252 ThrTPAAspGlnHnAlaProGlyAanTPAlaPheLysAspGlnValAlaAlaLeuSer 271
QY 631 TGGGTCCAGAAAGATGAGTTCTTCCGTGGGAGCCCGAGCTGTGACCATCTTTGGC 690
DB 272 TrpValGlnLysAanIleGlnPheGlyGlyAspProSerSerValThrIlePheGly 291
QY 691 GAGTCCGCGGAGCCATTAAGTGTCTTCTATCTATCTGTCTCCATGGCCAAAGCTTA 750
DB 292 GluSerAlaGlyAlaAlaIleSerValSerLeuIleLeuSerProMetAlaLysGlyLeu 311
QY 751 TTCCCAAAAGCCATGATGAGAGTGGGTGGGCATCATCCCTTACTGAGAGCCCATGAT 810
DB 312 PheHnIlyValAlaIleMeGlnSerIlyValAlaIleIleProTyrLeuGlnAlaHnAsp 331
QY 811 TATGAGAAAGTGAAGACTGTCAGGTGGTGGACATTTCTGTGTGAATGCGTCAGAC 870
DB 332 TyrGlnLysSerGluAspLeuGlnValAlaAlaHnSphCysGlyAanAanIleSerAsp 351
QY 871 TCTGAGGCGCTGCTGAGGTGCTGAGAACAAACCTTCAAGAGACTGTCAGCTCAGC 930
DB 352 SerGlnAlaLeuLeuArgCysLeuArgThrLysProSerIlyGlnLeuLeuThrLeuSer 371
QY 931 CAGAAACAAAGTCTTTCACTCGAGTGTGATGGTGTCTTCTTCCAAAGACCTCTA 990
DB 372 GlnLysThrLysSerPheThrArgValAlaAspGlyAlaPhePheProAspGlnProLeu 391
QY 991 GATGTTATGTTCTCAGAAAGCATTTAAAGCAATCTTCCATCATCGAGTCAATTAACAC 1050
DB 392 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAanAanHnIs 411
QY 1051 GAGTGTGGCTTCCGTGCTGCTTATGAAGAGGCTCTGAGATCTCGAGTGGCTCAACAG 1110
DB 412 GluLysGlyPheLeuLeuProMetLysGlnAlaProGlnIleLeuSerGlySerAanLys 431
QY 1111 TCCCTTGCCCTCATGATGATCAAAACATCTTGACATCCCGGCTCGATTTTGGACCTT 1170
DB 432 SerLeuAlaLeuHnIleLeuIleGlnAanIleLeuHnIleProProGlnLyrLeuHnIsLeu 451
QY 1171 GTGGCTATGATATCTTCCATGACAAAGCACTCCCTGATCGAATCCGAGACTTTCTG 1230
DB 452 ValAlaAanGlnLyrPheHnIleAspLysAanIleSerLeuThrGlnIleArgAspSerLeuLeu 471
QY 1231 GACTGTGTTGAGATGTGTCTTGTGTGCTCCGCTGATGATCAGAGTCAATATCAGACA 1290
DB 472 AspLeuLeuLeuLysAspValPhePheValValProAlaLeuIleThrAlaAspGlyrHnIsArg 491
QY 1291 GATGCTGTGTCACCTGTCTATCTTATGATTTTCGACACGCGCTCAGTCTTGAAGAC 1350
DB 492 AspAlaGlyAlaProValTyrPheTyrGluPheArgHnIsArgProGlnLysPheGlnAsp 511
QY 1351 ACAGAGCGGCTTTTGTCAAGCGACACGCTGATGAAGTCCGCTTGTGTTGGTGTGT 1410
DB 512 ThrLysProAlaPheValLysValAspHnIsAlaAspGlnValArgPheValPheGlyGly 531

QY 1411 GCCTTCTGAAGGGGAGATTTGTTATGTTGAAAGAGCCACGAGAGAGAGATTACTG 1470
DB 532 AlaPheLeuLysGlyLysAspIleValMetPheGlnGlyAlaThrGlnGlnGlnLysLeuLeu 551
QY 1471 AGCCGGAAGATGAGAAATTACTGGGCTACTCTTGTGTCGAACCGGGGAATCCATAAGGGAGAC 1530
DB 552 SerArgLysMetMetLysTyrTTPAlaThrPheAlaAanGlnArgLysAanProAanGlyAan 571
QY 1531 GACTGTCTCTGTGGCCAGCTTATTAATGACTGAGGAGTACCTTCAGCTGAGTCTTAAC 1590
DB 572 AspLeuSerLeuTyrTProAlaTyrAanLeuThrGlnGlnTyrLeuGlnLeuAspLeuAan 591
QY 1591 ATGAGCCCTCGGACGAGACTCTAAAGAACCGGGGGTGGATTTTGGACACGACCATGCC 1650
DB 592 MetSerLeuGlyGlnArgLeuLysGlnProArgValAspPheTyrPheThrIlePro 611
QY 1651 CTGATCTGTGCTGCTCGACATGCTCCACAGTCTCTCTTCTTCTTAACCTTCTCTCT 1710
DB 612 LeuIleLeuSerAlaAspAspMetLeuHnIsSerProLeuSerSerLeuThrPheLeuSer 631
QY 1711 CTCCTCCAGCCTTTCTTTTCTTGTGCTGCT 1743
DB 632 LeuLeuGlnProPhePhePheCysAlaPro 642
RESULT 6
ID ABP61005
XX ABP61005 standard; protein: 581 AA.
AC
XX ABP61005;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human protein. SEQ ID 92.
XX
KW Human, cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiac; antidiabetic; antitumor; antitubercular;
KW cerebroprotective; anorectic; metabolic; vaccinal; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW angioneurosis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
OS Homo sapiens.
XX
XX
PN NC0200250105-A1.
XX
PD 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WC-US049232.
XX
XX
PR 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX
PA (SMK) SMITHKLINE BEECHAM CORP.
PA (SMK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,

PI Mattensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX WPI; 2002-508784/54.
DR N-PSDS; ABO86170.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder
XX
PS Claim 1(a); Page 313-314; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnery, antiarteriosclerotic, antiparkinsonian, nootropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antitumor, vincriste, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metastatic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph nodes and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABP60965-ABP61019 represent novel human
CC proteins of the invention

SQ Sequence 581 AA;

Alignment Scores:

Pred. No.:	9.55e-242	Length:	581
Score:	2607.50	Matches:	499
Percent Similarity:	92.67%	Conservative:	7
Best Local Similarity:	91.39%	Mismatches:	25
Query Match:	82.26%	Indels:	15

US-10-023-515-3 (1-1746) X ABP61005 (1-581)

Qy	34	TGSGGCTTTTCCGATGTTCTCCAGCCGCCCTGTGGGACACAGAGTGGGAAAAACTGGG	93
		14 TriptalilerTriptalieuuLaalapro-----Thlygly	25
Qy	94	CCCTTCGTGAAGGGCCACAGAGAAACACAGCTGGATGATTCAAGGCAAGATC	153
Db	26	ProserHlaGlunGlyProGlnArganThrAgluengLYTPIleGlnGlyVal	45
Qy	154	ACTGTGCTGGGAAGCCCTGTGCCTGTGAACGTGTTCTCGAGTCCCTTGTGCTGCC	213
Db	46	ThrtVallengLysertProvalProvalaenValPheleugLysValProPhealLaalapro	65
Qy	214	CCGCTGGGATCCCTGGGATTTACGAACCCGAGCCTGCATGGCCCTGGGATTAATTGGCA	273
Db	66	ProleugLysertleuArgPheThrAnProGlnProAlaSerProTrrAPsAnleuArg	85
Qy	274	GAGGCACACCTCAACCTTAATTGTGGCCTCAGAACTCAGAGTGAGCTCTTAGATCAA	333
Db	86	GlulAlaThSerLysrrProanleuCySleugInanSerGluTrrPueuLeuAnpSln	105
Qy	334	CACATGCTCAAGGTGATTAACCGAATTCCGAGTGCAGAAAGCTGCCTTACCTGAC	393
Db	106	HismetleuLysValHisrrLysrrProLysPheGlyValSerGluAnpCySleuTrrLeuAn	125
Qy	394	ATCTAATGCGCTGCGCCACGCGCATACAGGCTCAAGCTCCCGCTTGTGTGTGCTCCA	453
		126 IleTrrYAlaProAlaHisAlaAspThrGlySerLysleuProValLeuValTrrPhePro	148

OY	454	GGAGTGCTTTCAAGACTGGCTCAGCCTCCATCTTTGATGGGTCCGCTGGCTCTAT	513
Db	146	GlYgYlAlPheLygYrNgYSerAlaSerIlePheApgGlySerAlaLeuAlaIaIaYr	165
OY	514	GAGAACGGCTGGGTGGTGGCTGACAGTCCGGGTAGAAATATTTGGTTCTTCACACA	573
Db	166	GlUnApValLeuValValValGlnTyArgLeuAlYlePheGlyPhePheThr	185
OY	574	TGGATCAGCATGTCTCCGGGAACTGGGCTTCAGAGACCAAGGTGCTGTCTCTGG	633
Db	186	TrpApgGlnHisAlaPProGlyAenThrAlaPheYAspGlnValAlaIaIaLeuSerTrp	205
OY	634	GTCCAGAAAGAACATCGAGTTCTTCGGTGGGAGACCCAGCTCTGTGACATCTTTGGCGAG	693
Db	206	ValGlnTyAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu	225
OY	694	TTCCGGGGAGCCAAATGTTTCTCAGCTTAATATGTCTCCATGGCCAAAGGTTATTC	753
Db	226	SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPhe	245
OY	754	CACAAACCCATCATGTGAGAGTGGGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGATAT	813
Db	246	HisYalAlaIleMetGluSerGlyValAlaIleIleProTyLeuGlnAlaHisAerTy	265
OY	814	GAGAAAGTAGAGACCTGCAAGTGGTGTGACATTTCTGTGZTAACAAATGCGTCAGACTT	873
Db	266	GluYsSerGlnAAsPLeuGlnValAlaIaHisPheCysGlnYAsnAlaIaSerAAsPser	285
OY	874	GAGGCCCTGCTGAGGTGCTCGAAGACAAACCTCCCAAGAGCTGTCGACCTTCAGCCAG	933
Db	286	GlUAlaLeuLeuArgCysLeuArgThrLysProSerIysGluLeuLeuThrLeuSerGln	305
OY	934	AAACAAAGCTTTCACTCGAGTGGTGTGAGTGGTCTTTCTTCTAATGAGCCCTGAT	993
Db	306	LysThrIysSerPheMetArgValValAAsPGLYAlaPhePheProAsnGluProLeuAAsP	325
OY	994	CTATTGTCTGAAAGCATTTAAGCAATTCCTTCATCATCGGAGTCAATTAACACAGAG	1053
Db	326	LeuLeuSerGlnYsAlaPheYsAlaIleProSerIleIleGlyValaAsnAlaIleGlu	345
OY	1054	TGTGGCTTCTGCTGCTGCTTAAGAGAG-----GCTCCTGAGATCCTCAGTGGCTCCAA	1107
Db	346	CysGlyPheLeuLeuPProMetValArgIleLeuAlaValHisThrAlaThrProSerAAsn	365
OY	1108	AAGGCCCTGGCCCTC-----CATCGTATACAAACATCCGACATCCCG	1152
Db	366	ArgAspAlaIaIaLeuAlaSerThrAlaGlyHisPheHisAArgArgHisGlnHisIlePro	385
OY	1153	CCTCAGATTTGGCACCCTTGGGTGTAATGAAATTACTCCATCAGACACATCCCTGACTGAA	1212
Db	386	ProGlnTyLeuHisIleLeuValAlaAsnGluTyPheHisAAsPArgHisIleSerLeuThrGlu	405
OY	1213	ATCCGAGACAGTCTTCTGACCTTGCTTGAAGATGTGTTCTTTGTGTCTCTGCACTGATC	1272
Db	406	IleArgAAsPLeuLeuAAsPLeuLeuGlyAAsPAlaPhePheValValProAlaLeuIle	425
OY	1273	ACAGCTGGAATATCACAGAGATGCTGTGGACCTGTCACTTCATAGATTTGGACACGG	1332
Db	426	ThrIaArgTyTrHisAAsPAspAlaGlyAlaProValTyPheTyGluPheAArgHisAArg	445
OY	1333	CCTCAGAGCTTTGAAGACAGAAAGCCGGCTTTTGTCAAAACCCAGACCACTGATGAAGTC	1392
Db	446	ProGlnTySereGluAAsPThrTySProAlaPheValIlyAlaAAsPHisAlaAAsPGLuVal	465
OY	1393	CGCTTTGTGTTCCGTGGTGCCTTCTTGAAAGGGAGACATTGTATGTTTGAAGAGCCAG	1452
Db	466	ArgPheValaPheGlyGlyAlaPheLeuTyGlyAAsPAlaIaIaMetPheGluGlyAlaThr	485
OY	1453	GAGAGAGAAAGTTACTGAGACCCGGAAGATGATGAATATCTGGGCTACCTTTGCTCCAAAC	1512
Db	486	GlnGlnIuIuIySereLeuSerArgLysMetMetLysTyTrIpaAlaThrPheAlaAArgThr	505


```

QY 1174 GCTATGAAATCTTCATGACAAGCACTCCGTAAGTAATCCGAGACAGTCTTGAC 1233
DB 386 AAlaangluryrPhehlaAspyshehSerleuhtgullleayAspSerleuenuasp 405
QY 1234 TTGCTGGAGATGCTGTTTGTGTGCTCCCTGCACTGATCAGAGCTGATATCAGAGAT 1293
DB 406 LeuenuglayarpValPhePheValProAlaIleuIlethnAlaArgTyrHlaArg--- 424
QY 1294 GCTGTGACCTGCTTACTTCTATGAGTTTGCGACCGGCTCACTGCTTGAACAACAG 1353
DB 424 ----- 424
QY 1354 AAGCGGCTTTTGTCAAGCGACACCGCTGATGAAGTCCGCTTGTGTGCTGTGTC 1413
DB 424 ----- 424
QY 1414 TTCCTGAAGGGGACATTTGATGTTTCGAAGAGACCGACGAGAGAGAGATTACTGAC 1473
DB 425 -----GluGlYAlaThrGluGluGluLeuLeuSer 435
QY 1474 CGGAAGATGATGAATACTGGGCTACCTTTGCTGCAACCGGGAATCTTAATGGAGACAC 1533
DB 436 ArglysmetMetyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 455
QY 1534 CTGTCTGTGCGGCGCTTATATCTGACTGAGACTGCTCCGCTGCTGACTGAAACATG 1593
DB 456 LeuSerleuTrpProAlaTyrAsnleuThrGluGlnTyrleuGlnLeuAspSerleuMet 475
QY 1594 AGCTCGACAGAGACTCAAAAGAACCGGCGGTGATTTTGGACGACGACATCCCTCTG 1653
DB 476 SerleuGlyGlnArgleuLeuGluProArgValaIlePheTrpThrleuSerleu 495
QY 1654 ATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCTTACTTCTCTCTCTC 1713
DB 496 IleuSerAlaSerAspMetleuHisSerProleuSerleuThrPheleuSerleu 515
QY 1714 CTCACGCTTCTTCTTCTTCTTGTGCTCCT 1743
DB 516 LeuGlnProPhePhePheCysAlaPro 525

RESULT 8
ABUS4639 ID ABUS4639 standard; protein; 581 AA.
AC ABUS4639;
AC 03-JUN-2003 (first entry)
DT 03-JUN-2003
DE Human NOXV polypeptide #98.
XX Human NOXV polypeptide #98.
KW Human; NOXV; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
OS Homo sapiens.
PN WO200281498-A2.
XX 17-OCT-2002.
XX 03-APR-2002; 2002WO-US010780.
XX 03-APR-2001; 2001US-0281086P.
XX 03-APR-2001; 2001US-0281136P.
XX 05-APR-2001; 2001US-0281863P.
XX 05-APR-2001; 2001US-0281906P.
XX 06-APR-2001; 2001US-0282020P.

```

```

PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.

(CDRA-) CDRAGEN CORP.
XX Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
XX Patuxian M, Liu X, Gusev VY, Li L, Vernet CM, Zerkusen BD;
XX Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
XX Padigaru M, Shinkels RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W;
XX Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ;
XX MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
XX Ellerman K;
XX MPI: 2003-046858/04.
XX N-PDB; ABX72267.
XX New isolated NOXV polypeptide useful for treating atherosclerosis,
XX metabolic disorders, diabetes, obesity, infectious disease, anorexia,
XX neurodegenerative disorders, Alzheimer's disease and cancer.
XX Claim 1; Page 303; 666pp; English.
XX The invention relates to human polypeptides, termed NOXV, and the
XX polynucleotides encoding them. The polypeptides and polynucleotides are
XX useful for diagnosing disease, and screening for potential therapeutic
XX agents. The sequences are useful for treating metabolic disorders,
XX cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
XX stenosis, atrial septal defect (ASD), atrioventricular canal defect,
XX ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
XX septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
XX atherosclerosis, obesity, infectious diseases, anorexia, neurodegenerative
XX disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
XX haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
XX and cancer. Sequences ABUS4542-ABUS4647 represent human NOXV polypeptides
XX of the invention
XX SQ Sequence 581 AA;
XX
XX Alignment Scores:
XX Pred. No.: 8.79e-241 Length: 581
XX Score: 2597.50 Matches: 497
XX Percent Similarity: 95.07% Conservative: 4
XX Best Local Similarity: 94.31% Mismatches: 17

```

Query Match: 81.94% Indels: 9
 DB: 6 Gaps: 3
 us-10-023-515-3 (1-1746) x ABUS4639 (1-581)

```

QY 91 GGGCTTGTCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 150
DB 27 GTPProSerAlaGluGlyProGlnArgAenThrArgLeuGlyTyrIleGlnIlySerGln 46
QY 151 GTACATGTGCTGGGAAGCCCTGTGCTGGAACGTGTCTCTGAGAGTCCCTTGTGCT 210
DB 47 ValThrValLeuGlySerProValProValAanValPheLeuGlyValProPheAlaIa 66
QY 211 CCCCCTGGGATCCCTGGATTTTACGAACCCGAGCCCTGCATCCCTGGGATTAATCTG 270
DB 67 ProProLeuGlySerLeuArgPheThrAenProGlnProAlaSerProThrAenPhe 86
QY 271 CGAGAAAGCCACTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAGT 330
DB 87 ArgGlnAlaThrSerTyrProAenLeuGlyLeuGlnAanSerGlnThrLeuLeuAenAsp 106
QY 331 CAACACATGCTCAAGTGTGATTAACCCGAATTCGAGAGTGTCAAGAACTGCTCTTACTG 390
DB 107 GlnHisMetLeuValHisValHisTyrProLysPheGlyValSerGlnAspCysLeuTyrLeu 126
QY 391 AACATCTATGGCGCTGCCACGCGCGATACAGGCTCCAGAGCTCCCGCTGTGGTGTGCTTC 450
DB 127 AsnIleTyrAlaProAlaHisAlaAspThrGlySerTyrLeuProValLeuValThrPhe 146
QY 451 CCAGAGAGTGGCTTCAAGACTGTGCTCAGCTTCATCTTTGATGGTCCGCTGGCTGGCC 510
DB 147 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAenGlySerAlaLeuAlaIa 166
QY 511 TATGAGACGTGCTGTGATGCTGCTCAGTACCGGCTAAGGAATTTGGTTTCTTCAAC 570
DB 167 TyrGlnAspValIleValValValGlnTyrHisLeuGlyIlePheGlyPhePheThr 186
QY 571 ACATGGGATCAGATGCTCCGCGGAACTGGGCTTCAAGAACAGAGTGGCTGTCTGCC 630
DB 187 ThrTyrAenGlnHisAlaPheGlyAsnThrAlaPheLysAenGlnValAlaIleAenSer 206
QY 631 TGGGTCCAGAAAGAACATGAGATTCTTCGCTGGGAACTCCAGCTCTGTGACCATTTTGGC 690
DB 207 TyrValGlnTyrAanIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly 226
QY 691 GAGTCCCGCGGAGCATTAAGTGTCTTACTTACTGCTCCATGGCCAAAGGCTTA 750
DB 227 GluSerIleGlyAlaIleSerValSerSerLeuLeuSerProMetAlaLysGlyLeu 246
QY 751 TTCCCAAAAGCCATCATGAGAGTGGGCTGGCCATCATCCCTTACCTGAGGCCCATGAT 810
DB 247 PheHisLysValIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 266
QY 811 TATGAGAAAGTGAAGACCTGCAGAGTGTGCAATTTCTGTGTAACATGCTGTCAAC 870
DB 267 TyrGlnLysSerGlnAspLeuGlnValValAlaHisPheCysGlyAanAanAlaSerAsp 286
QY 871 TCTGAGGCGCTGTGAGTGTGCTGAGAGCAAAACCTCCAGAGAGTGTGAGCCCTGAGC 930
DB 287 SerGlnAlaLeuLeuArgCysLeuArgThrLysProSerTyrGlnLeuLeuThrLeuSer 306
QY 931 CAGAAAACAAAGTCTTCACTCGAGTGTGATGAGTCTTTCTTCTTAATGAGCCCTTA 990
DB 307 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAenGlnProLeu 326
QY 991 GATCTATGTCTCAGAAAGCATTTAAAGCAATTCCTCCATCATCGAGAGCATTAACAC 1050
DB 327 AspLeuLeuSerGlnTyrAlaPheLysAlaIleProSerIleIleGlyValAanAanHis 346
QY 1051 GAGTGTGCTTCTGTGCTCATGAGAGAG-----GCTCTGAGAGATCCAGTGGCTCC 1104
DB 347 GluCysGlyPheLeuLeuProMetValArgIleLeuAlaValHisThrAlaThrProSer 366

```

```

QY 1105 AACAGATCCCTTGGCCCTC-----CATCTGATACAAACATCTTGCAATC 1149
DB 367 AsnArgAspAlaValLeuAlaSerThrAlaGlyHisPheHisArgArg-----HisIle 384
QY 1150 CCGGCTCAGTATTTTGGACCTTGAGCTTAATGAAATCTCCATGCAAGACATCCCTGAT 1209
DB 385 ProGlnIleTyrLeuHisLeuValAlaAenGlnTyrPheHisAspLysHisSerLeuThr 404
QY 1210 GAAATCCGAGACAGTCTTTTGAAGTCTTGGACTTGTGAGATGTGTTCTTTGTGCTCCGCACTG 1269
DB 405 GluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValProAlaLeu 424
QY 1270 ATCAGAGTGCATTTCAAGAGAGTGTGCTGACCTGTCTTACTTCTTATGAGTTTGGGAC 1329
DB 425 IleThrAlaArgTyrHisArgAspAlaGlyAlaLeuProValTyrPheTyrGlnPheArgHis 444
QY 1330 CGGCTCAGTGTCTTGAAGACACAGAGCCGGCTTTGTCAAAAGCCGACACGCTGATGA 1389
DB 445 ArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHisAlaAspGln 464
QY 1390 GTCCGCTTTGTGTGTGCTGTGCTTCTGAAAGGGGACATGTTATGTTTGAAGAGCC 1449
DB 465 ValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlnGlyAla 484
QY 1450 ACGGAGAGAGAGAGTACTGAGCCGGAAGATGAATAATCGGCTTACTTGTCTGCA 1509
DB 485 ThrGlnGlnGlnLysLeuLeuSerArgLysMetLysTyrThrAlaThrPheAlaArg 504
QY 1510 ACCGGAAATCCTAATGAGGAACGACTGTCTGTGGCCAGCTTATATCTGACTGAGCAG 1569
DB 505 ThrGlyAanProAenGlyAanAspLeuSerLeuThrProAlaTyrAenLeuThrGlnGln 524
QY 1570 TACTCTCAGCTGGAATTTGAACATGAGCTCTGACAGAGATCTCAAGAACCCGGGTGAT 1629
DB 525 TyrLeuGlnLeuAspLeuAenMetSerLeuGlyGlnArgLeuLysGlnProArgAsp 544
QY 1630 TTTTGACACAGACCATCCCC 1650
DB 545 ValThrValThrGlyTyrPro 551

RESULT 9
ADR19663
ID ADR19663 standard; protein; 618 AA.
XX
AC ADR19663;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human drug metabolising enzyme (DME)-2 protein sequence.
XX
KW drug metabolising enzyme; DME; cytostatic; immunosuppressive;
KW antiinflammatory; endocrine; ophthalmological; gastrointestinal;
KW hepatocarcinoma; cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; endocrine disorder; eye disorder;
KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
human.
XX
OS Homo sapiens.
XX
PN MO200226988-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US030662.
XX
PR 29-SEP-2000; 2000US-0236947P.
PR 06-OCT-2000; 2000US-0238864P.
PR 20-OCT-2000; 2000US-0242323P.
PR 09-NOV-2000; 2000US-0247581P.
PR 16-NOV-2000; 2000US-0249519P.
PR 22-NOV-2000; 2000US-0252834P.
PR 30-NOV-2000; 2000US-0250567P.
XX

```

PA (INCY-) INCYTE GENOMICS INC.

XX Azimzai Y, Baughin MR, Boroweky ML, Ding L, Dugan BW,

PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA,

PI Lai P, Lee EA, Lu DM, Nguyen DB, Arvizu C, Policky JL, Ramkumar J,

PI Ring HZ, Sarjamaala MS, Tang YT, Tribouley CM, Narinder WK,

PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H,

DR WPI: 2002-362498/39.

DR N-PSDB: ADR19681.

XX Human drug metabolizing enzymes, useful in the diagnosis and treatment of

PT disorders associated with aberrant (DME) activity, e.g., cancer and

PT autoimmune disorders.

XX Claim 1; SEQ ID NO 2; 142pp; English.

CC This invention relates to novel drug metabolizing enzymes (DME) and the
CC nucleotide sequences which encode them. The invention may be useful for
CC the development of compounds with a cytostatic, immunosuppressive,
CC antiinflammatory, endocrine, ophthalmological, gastrointestinal or
CC hepatotropic activity acting as an agonist or antagonist of drug
CC metabolizing enzyme activity. The invention may be used in the diagnosis
CC and treatment of disorders associated with decreased or increased
CC expression or activity of drug metabolizing enzymes. Such disorders
CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
CC endocrine, eye, gastrointestinal (including liver disorders) and
CC metabolic disorders. The present sequence is that of a human drug
CC metabolizing enzyme (DME) of the invention. Note: This sequence did not
CC form part of the printed specification but was obtained in electronic
CC format from EPO.

XX Sequence 618 AA;

Alignment Scores:

Pred. No.: 6,71e-240 Length: 618

Score: 2588.50 Matches: 491

Percent Similarity: 94.62% Conservative: 1

Best Local Similarity: 94.42% Mismatches: 5

Query Match: 81.66% Indels: 23

DB: Gaps: 1

US-10-023-515-3 (1-1746) x ADR19663 (1-618)

QY 91 GGGCTTCTGCTGAAGGCGCAAGAGAACACAGGCTGGATTCAGGCGCAAGCA 150

DB 92 G1YProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrIleGlnIyLysGln 111

QY 151 GTCACGTGCTGGGAAGCCGTGCTGTGAACGTTCCTCGAGTCCCTTGCTGCTGCT 210

DB 112 ValThrValIleuGlySerProValProValAsnValPheLeuGlyValProPheAlaIa 131

QY 211 CCCCCGTGGATCCCTGCAATTAACAACCCGACGCTGATCGCCCTGGATTAATTG 270

DB 132 ProProIleuGlySerIleuArgPheThrAsnProGlnProAlaSerProTyrAspAsnIeu 151

QY 271 CGAAGACCACTCTCACTCAATTTGCTGCTCGAATCTCAAGTGGCTGCTTGAAT 330

DB 152 ArgGluAlaThrSerTyrProAsnIeuCysLeuGlnAsnSerGluTyrPheIleuAsp 171

QY 331 CAACACATGCTCAAGGTGCATTAACCGAAATTCGAGGTGCAGAAAGATGCTTACCTG 390

DB 172 GlnIleMetLeuLysValIleTyrProLysPheGlyValSerGluAspCysLeuTyrIeu 191

QY 391 AACATCTATGCGCTGCCACGCGCATACAGGCTTCAAGCTCCCGCTTGGTGTGCTTC 450

DB 192 AsnIleTyrAlaProAlaIleAlaAspThrGlySerTyrIleuProValIleuValIlePhe 211

QY 451 CGAAGAGTGCCTTCAAGACTGGCTCAGCTTCATCTTGAATGGTCCGCTGGCTGCC 510

DB 212 ProGlyGlyAlaIlePheLysThrGlySerAlaSerIlePheAspGlySerAlaIleuAlaIa 231

QY 511 TATGAGACGTCGCTGCTGTGCTGCTCAAGTACCGGCTAAGAAATTTGGTTTCTTACC 570

DB 232 TyrGluSerValIleuValIaValIaGlnTyrArgLeuGlyIlePheGlyPheSerThr 251

QY 571 ACATGGATTCAGCATGCTCCCGGGGAACCTGGCCCTTCAAGACCAAGTGGCTGCTGCC 630

DB 252 ThrTyrAspGlnIleAlaProGlyAsnTyrAlaPheLysAspGlnValAlaIleuSer 271

QY 631 TGGGTCCAGAAAGATCATGAGTTCCTCGATGGGAGCCCACTGTCAGCATCTTGGC 690

DB 272 TrpValGlnLysAsnIleGluPhePheGlyIleAspProSerSerValThrIlePheGly 291

QY 691 GAGTCCGGGAGGACCATAGTGTTCCTTACTGTTACTGTTCCCATGGCCAAAGCTTA 750

DB 292 GluSerAlaGlyAlaIleSerValSerSerIleuSerProMetAlaIyGlyIeu 311

QY 751 TTCACAAAGCCATCATGAGAGAGTGGGTGGCCATCATCCTTACCTGAGAGCCCATGAT 810

DB 312 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaIleAsp 331

QY 811 TATGAGAAAGATGAGACCTGAGGTGGTTCACATTTCTGTGGTAAATGCTCAGAC 870

DB 332 TyrGluLysSerGluAspLeuGlnValValAlaIlePheCysGlyAsnAsnIleSerAsp 351

QY 871 TCTGAGGCTCTGCTGAGGTGCTGAGACAAACCTTCCAAAGAGCTGTGACCTCAGC 930

DB 352 SerGluAlaIleuLeuArgCysLeuArgThrLysProSerLysGluLeuThrLeuSer 371

QY 931 CAGAAACAAAGCTCTTTCACCTGAGGTGTGATGCTGCTTCTTCCTAAGACCTTCA 990

DB 372 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProIeu 391

QY 991 GATCATTTGCTCGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACCA 1050

DB 392 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHis 411

QY 1051 GAGTGTGCTTCTGCTGCTGCTATGAAGAGGCTCCTAGATCCTCAGTGGCTCAACAG 1110

DB 412 GluCysGlyPheLeuLeuProMet----- 419

QY 1111 TCCCTTGGCTTCATCTGATCAAAACATCCTGCACATCCCGCTCAGATTTTGACCTT 1170

DB 420 -----HisIleProProGlnTyrIleuHisIleu 428

QY 1171 GTGGCTATGAAATCTTCCATGACAGACCTCCGACTGAATCCGAGACAGTCTTCTG 1230

DB 429 ValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleAspSerIleuLeu 448

QY 1231 GACTTGTGAGATGATGTTCTTGTGTGCTCCCTGCACTGATCAGAGCTGATATCACAGA 1290

DB 449 AspLeuLeuGlyAspValPhePheValAlaProAlaLeuIleThrAlaArgTyrHisArg 468

QY 1291 GATCTGTGTCACCTGTCTACTTCTATGAGTTTGGCAACCGGCTCAGTCTTGAAGAC 1350

DB 469 AspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAsp 488

QY 1351 ACGAAGCGGCTTTTTCGAACCGGACCAAGCTGATGAAGTCCGCTTGTGGTGGT 1410

DB 489 ThrLysProAlaPheValLysAlaAspHisValAlaAspGluValArgPheValPheGlyGly 508

QY 1411 GCCTTCTGAGAGGGGACATGTTATGTTTCGAAAGAGCCAGGAGAGAGAAAGTTACTG 1470

DB 509 AlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluLysLeuLeu 528

QY 1471 AGCCGAAAGATGATGAATACTAGGCTTACCTTGTCTGACACCGGGAATCTTAATGGGAAC 1530

DB 529 SerArgLysMetCysLysTyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 548

QY 1531 GACCTGTCTGTGTGGCCAGCTTATATCTGACTGAGAGTACCTCCAGCTGGACTGTAAC 1590

DB 549 AspLeuSerLeuTyrProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsn 568

QY 1591 ATGAGCTTCGAGACAGATCAAAAGAACCGCGGCTGATTTTGGACCAACCATCCCC 1650

DB 569 MetSerLeuGlyGlnArgLeuYsgIuProArgArgAspValTrpValThrGlyTyrrPro 568
RESUL T 10
ADB64065
ID ADB64065 standard; protein; 469 AA.
XX ADB64065;
AC
XX
DT 04-DEC-2003 (first entry)
DE Human protein encoded by clone BRAMH20021910.
XX
XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX
XX Homo sapiens.
OS
XX
XX EPI308459-A2.
XX
XX 07-MAY-2003.
XX
XX 28-MAR-2002; 2002EP-00007401.
XX
XX 05-NOV-2001; 2001JP-00379238.
XX
XX 25-JAN-2002; 2002US-00350978.
XX
XX
XX (HELI-) HELIX RES INST.
XX (REMS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
XX MPI: 2003-450961/43.
XX
XX N-PSDB; ADB62095.
XX
XX
XX New polynucleotides and polypeptides, useful for developing a diagnostic
XX marker or medicines for regulation of their expression and activity, or
XX as targets of gene therapy.
XX
XX
XX Claim 1; Page: 222pp; English.
XX
XX The invention discloses a polynucleotide comprising a sequence selected
XX from 1970 fully defined nucleotide sequences which encode novel
XX polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX or its partial peptide, an antibody binding to the polypeptide or peptide
XX of the polynucleotide, immunologically assaying the polypeptide or
XX peptide of the polynucleotide by contacting the polypeptide or peptide
XX with the antibody of the encoded protein, and observing the binding
XX between the two, a transformant carrying the polynucleotide in an
XX expressible manner and an antisense polynucleotide. The oligonucleotide
XX is useful as a primer for synthesizing the polynucleotide, or as a probe
XX for detecting the polynucleotide. The polynucleotides and encoded
XX proteins are useful as pharmaceutical agents and many disease-related
XX genes may be included in them, for developing a diagnostic marker or
XX medicines for regulation of their expression and activity, or as targets
XX of gene therapy. The genes are involved in tissue and/or cell
XX regeneration. Membrane proteins, signal transduction-related proteins,
XX transcription-related proteins, disease-related proteins and genes
XX encoding them can be used as indicators for diseases (e.g. osteoporosis,
XX neurological diseases, cancer, tumours. The cDNA may be used to regulate
XX the activity or expression of the encoded protein to treat diseases. The
XX sequence presented is a protein of the encoded protein. Note: Some of the
XX sequence data for this patent is not represented in the printed
XX specification, but is based on sequence information supplied by the
XX European Patent Office.
XX
XX Sequence 469 AA;
SQ
XX
XX Alignment Scores: 5.42e-227 Length: 469
XX
XX Pred. No.: 5.42e-227 Length: 469

Score: 2454.00 Matches: 467
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.57% Mismatches: 0
Query Match: 77.41% Indels: 0
DB: 7 Gaps: 0
US-10-023-515-3 (1-1746) x ADB64065 (1-469)
QY 337 ATGCTCAAGGTGATACCCGAAATTCGAGTGTCAAGACCTGCTTACCTGAACATC 396
DB 1 MetLeuYsValIhIstYrProLyAspHeGlyAlaSerGluAspCySleuTyLeuAsnIle 20
QY 397 TATGGCTGCGCCAGCGCCATACAGGCTCCCAAGTCCCGCTTGGTGTGTTCCAGGA 456
DB 21 TyrAlaProAlaIhIstYrProLyAspHeGlyAlaSerGluAspCySleuTyLeuAsnIle 40
QY 457 GGTGCTTCAAGACTGCTGAGCTCCATGCTTGTATGGTCCGCTGCTGCTTATGAG 516
DB 41 GlyAlaPheYsThrGlySerAlaSerIlePheAspIleSerAlaLeuAlaIstYrGlu 60
QY 517 GACGTGCTGTTGTGTGCTCAGTACCGGCTAGAAATATTGGTTCTTACCACATGG 576
DB 61 AspValLeuValIstYrProLyAspHeGlyAlaSerGluAspCySleuTyLeuAsnIle 80
QY 577 GATCAGCATGCTCCGCGGAACTGGGCTTCAAGGACCAAGTGGCTGCTGCTGCTGCTC 636
DB 81 AspGlnIhIstYrProLyAspHeGlyAlaSerGluAspCySleuTyLeuAsnIle 100
QY 637 CAGAAAGACATGAGTCTTGTGTTGGGACCCGAGCTGCTGACATCTTGGCGAGTCC 696
DB 101 GlnYsAsnIleGlyPhePheGlyAlaSerProSerSerAlaThrIlePheGlyGluSer 120
QY 697 GCGGAGCCATTAAGTGTCTTACTGTTATCTGCTTCCATGCGCAAGGCTTATTCAC 756
DB 121 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaYsgIleuPheIhIst 140
QY 757 AAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGAGGCGCATGATTGAG 816
DB 141 LysAlaIleMetGluSerGlyValAlaIleIleProTyLeuGluAlaIstYrGlu 160
QY 817 AAGATGAGGACCTGACAGTGTGTGACATTTGTGTGTAACATGCTGAGCTGAG 876
DB 161 LysSerGluAspLeuGlnValIstYrProSerSerAlaIstYrGluAspCySleuTyLeuAsnIle 180
QY 877 GCGCTGTGAGGCTGCGGAGCAAAACCTCCAAAGAGTGTGACCTGACCGCAAGAA 936
DB 181 AlaLeuIleuArgYsLeuIstYrProSerSerAlaIstYrGluLeuIstYrGlu 200
QY 937 ACAAGTCTTCACTGAGTGTGATGAGTGTCTTCTTATGAGCTTATGAGCTTGA 996
DB 201 ThrIleSerPheThrArgValAlaAspGlyAlaPhePheProAsnIleuPheAspLeu 220
QY 997 TTGTCTCAAAAGCATTTAAAGCAATTCCTTCATGATGAGTCAATGACAGAGTGT 1056
DB 221 LeuSerGlnIstYrProSerSerAlaIstYrGluLeuIstYrGlu 240
QY 1057 GGTTCCTGCTGCTTGAAGAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1116
DB 241 GlyPheLeuIleuProMetGlyValAlaProGlnIleLeuSerGlySerAsnIstYrGlu 260
QY 1117 GCCCTCATCTGATACAAACATCTGACATCCGCTCAGTATTTGGACCTTGTGCT 1176
DB 261 AlaLeuIleuIleGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 280
QY 1177 AATGAATATCTTCATGACAAAGCATCTGCTGATGATGATGATGATGATGATGATG 1236
DB 281 AsnGlnIstYrPheIstYrPheIstYrPheIstYrPheIstYrPheIstYrPheIstYr 300
QY 1237 CTGGAGATGTGTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1296
DB 301 LeuGlyAspValPhePheValIstYrProAlaLeuIleuIleuIleuIleuIleuIleuIleu 320
QY 1297 GGTGACCTGTCTACTTATGAGTGTTCGAGCAGGCTCAGTGTCTTGAAGACAGAA 1356

```

Db      321 GlyAlaProValIYrPheTcGluPheArghIaYgProGlnCybPheGluPheThrIys 340
Qy      1357 CCGGCTTTTTCGAAGCCGACCCATGTAAGTCCTGTTGTTGCTGGTGTGCTTC 1416
Db      341 ProAlaPheValIYsAlaAspHisAlaAspGluValArgPheValPheGlyGlyAlaPhe 360
Qy      1417 CTGAAGGGGAGCATGTTATGTTCTGAAGAGCCAGCAGAGAGAGAGATTACTGAGCCGG 1476
Db      361 LeuIysGlyAlaPheIleValMetPheGluGlyAlaThrGluGluIleuLeuSerArg 380
Qy      1477 AAGATGATGAATACTGAGGCTTACCTTTGCTCGAACCCGGAATCTTAATGGAGCACTTG 1536
Db      381 LysMetMetLysIYrTrrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeu 400
Qy      1537 TCTCTGTGGCCAGCTTATATATGACTGAGCAGTACTCTTCAGCTGACTTGAACATGAGC 1596
Db      401 SerLeuTrpProAlaIYrAsnLeuThrGluGlnIYrLeuGlnLeuAspLeuAsnMetSer 420
Qy      1597 CTCGAGCAGAGACTCAAGAACCAGCGGGTGATTTTGGACAGACCATCCCTGATC 1656
Db      421 LeuGlyGlnAlaGlyLeuGluProArgValGluPheTrpMetSerThrIleProLeuIle 440
Qy      1657 CTGTCTGCTCCGACATCCTCCACAGTCCTCTTTCTTAACTTCTCTCTCTCTC 1716
Db      441 LeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeuLeu 460
Qy      1717 CAGCCTTTCTTTTCTTTTGTGCTCT 1743
Db      461 GlnProPhePhePheCysAlaPro 469

```

RESULT 11

ADFS0145
ID ADFS0145 standard; protein; 542 AA.

```

XX AC ADFS0145;
XX DT 12-FEB-2004 (first entry)
XX DE Cat cauxin protein SEQ ID NO:2.
XX KW cat; cauxin; cat kidney disease marker; kidney disease.
XX OS Felis catus.
XX FN JP2003250575-A.
XX PD 09-SEP-2003.
XX PE 04-MAR-2002; 2002JP-00057908.
XX PR 04-MAR-2002; 2002JP-00057908.
XX PA (TOHO-) TOHOKU TECHNOARCH KK.
XX DR MPI: 2004-002277/01.
XX N-PSDB: ADFS0144.
XX PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
XX PS Claim 3; SEQ ID NO 2; 33pp; Japanese.
XX
XX The present sequence represents a cat cauxin protein (I) or its salt,
XX which is cat kidney disease marker. Also described: (i) a partial peptide
XX (ii) of (i); (2) a nucleic acid (iii) encoding (i) or (ii); (3) a vector
XX (iv) comprising (iii); (4) a transformed host (v) comprising (iii) or
XX (iv); (5) producing (i) or (ii) by culturing (v); (6) an antibody (vi)
XX which couples specifically with (i) or (ii); (7) diagnosing cat kidney
XX disease which involves measuring (i) quantitatively, and where reduction
XX of amount of (i) indicates presence of the disease; (8) a cat kidney
XX disease diagnostic agent comprising (i) labelling agent, a reagent which
XX measures the biological activity of urinary (i) or (vi); and (9) a cauxin

```

CC detection kit which measures cauxin in a test sample. (i) is useful as a
 CC cat kidney disease marker and (vi) is useful for diagnosing cat kidney
 CC disease. (ii) enables detection of cat kidney disease simply and
 CC correctly. (i) provides an early marker for the disease, and replaces
 CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
 CC blood testing.

XX SQ Sequence 542 AA;

Alignment Scores:

Pred. No.:	Length:
1e-181	542
Score:	1985.00
Percent Similarity:	83.27%
Best Local Similarity:	71.73%
Query Match:	62.62%
DB:	Gaps: 2

US-10-023-515-3 (1-1746) x ADFS0145 (1-542)

```

Qy      91 GGGCTTCTGCTGAAGGGCCACAGAGAACAGAGCTGGAGTTCAGGGCAAGCA 150
Db      22 GlyProAlaAlaAspAlaProValArgSerThrArgLeuGlyTrrValArgGlySerGln 41
Qy      151 GTCACTGTCTGGAGAGCCCTGTGCTGTGAACGTTTCTCGAGTCCCTTGTCTGCT 210
Db      42 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTrrAlaAla 61
Qy      211 CCCCCTGGGATCCCTCGATTACGAACCCGACGCTGCATCGCCCTGGATACCTTG 270
Db      42 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTrrAlaAla 61
Qy      62 ProProLeuGlyProLeuAlaPheLeuSerGlnProValAlaLeuProGlyAsnAspPhe 81
Db      271 CGAGAGCCACCTCCTTACCTTAATTTGTGCTCCGACGACTCAGAGTGGCTGCTTGAAT 330
Qy      82 ArgAsnAlaTrpSerTrrProLysLeuCybPheGlnMetLeuGluTrrPheValSerTrr 101
Db      331 CAACATGCTCAAGGTGCTATTAACCCGAATTCGAGTGTGAGAGACTGCTTACCTG 390
Qy      102 GlnHisValLeuLysValArgTrrProLysLeuGlnAlaSerGlnAspCysLeuTrrLeu 121
Qy      391 AACATCTATGCGCTGCCCGACCGGATACAGGCTTCAGGCTCCCGCTTGGTGTGCTTC 450
Db      122 AsnIleTrrAlaProAlaHisAlaAspAsnGlySerAsnLeuProValMetValTrrPhe 141
Qy      451 CCAGAGGTGCTTCAAGACTGCGTCAAGCTCCATCTTGAATGGGTCCGCTGGCTGCC 510
Db      142 ProGlyGlyAlaPheLysMetGlySerAlaSerPheAspGlySerAlaLeuAlaAla 161
Qy      511 TATGAGACGTGCTGTTGTGCTGCTCCAGTACCGGCTAGGAAATTTGTTCTTACC 570
Db      162 TyrGluAspValLeuIleValThrThrIleTrrArgLeuGlyIlePheGlyPhePheAsp 181
Qy      571 ACATGGATCAGCATGCTCCGAGAACTGGGCTTCAAGAACAGAGTGGCTGCTGCC 630
Db      182 ThrGlyAspGlnHisAlaArgIYrAsnTrrPalLeuLeuAspGlnValAlaLeuThr 201
Qy      631 TGGTCCAGAGAACATGAGTTCCTGCTGGGAGCCCACTGTCATCATCTTGGC 690
Db      202 TrrValArgAspAsnIleGlnPhePheGlyGlyAspProArgSerValThrIlePheGly 221
Qy      691 GAGTCCGGGAGGACATGATGTTTCTTACTTACTGTCTCCATGCGCAAGGCTTA 750
Db      222 GluSerAlaGlyAlaIleSerValSerSerLeuLeuSerProIleAlaAsnGlyLeu 241
Qy      751 TTCACAAAGCCATCATGAGAGAGTGGGTCATCATCTTACTG--GAGGCCAT 807
Db      242 PheHisLysAlaIleMetGluSerGlyValAlaIleLeuProLeuMetArgProPro 261
Qy      808 GATTATGAGAGATGAGAGCTCAGCTGCTGTCACATTTCTGTGAAACAAATGCGCA 867
Db      262 GlyAspGlnArgGlyLysAspLeuGlnValLeuAlaArgIleCybGlyCybHisIleAsp 281
Qy      868 GACTCTGAGGCTCTGCTGAGTGGCTGAGCAAAACCTTCAAGAGACTGACCTG 927

```



```

Db      282 AppSerAla1AlaLeuLeuGlnCysLeuArgAlaIyProSerGlnGluLeuMetAspIle 301
Qy      928 AGCCAGAAAACAAGCTTTCACCTGACGGTGGATGAGCGCTTTCTTTCTTAATGAGCCT 987
Db      302 SerIySylLeuThrPheSerIleProValIleAspAspPhePheProAspGluPro 321
Qy      988 CTAGATCTAATGCTCAGAAAGCATTTAAAGCAATCCCTTCATCATCTGAGATCAATAAC 1047
Db      322 ValAlaLeuLeuThrGlnIySylAlaPheAsnSerValProSerIleIleGlyAlaAsn 341
Qy      1048 CACGAGTGGGCTTCTGCTGCTCCTATGAAAGAGGCTCTGAGATCTCATAGGCTTCAAC 1107
Db      342 HisGlnCysAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyGlySerAsn 360
Qy      1108 AAGTCCCTTGGCCCTCATGTATACAAAACATCCGACATCCCGCTCAGATTTGGAC 1167
Db      361 ArgSerLeuAlaLeuIyLeuValHisThrPheLeuAsnIleProThrGlnIyLeuHis 380
Qy      1168 CTTCGCTTAATGAACTATCTTCATGACACAGCACTCCCTGACTGAATCCGAGACAGCTTT 1227
Db      381 LeuValAlaAspHisIeTyPheTyAsnIyHisSerProValGluIleArgAspSerPhe 400
Qy      1228 CTGAGCTTGGCTTGGAGATGTGTCTTTGTGCTCCTGCACTGATCAAGCTGATATAC 1287
Db      401 LeuSerPheLeuGlnIyAspValLeuPheValValProGlnIyValThrAlaArgTyHis 420
Qy      1288 AGAGATGCTGGTGACCTGTCTACTTCTATGATTCGGACCGGCTCAGTGCTTGA 1347
Db      421 ArgAspAlaGlnIyAlaProValIyPheTyGlnPheGlnHisProGlnCysLeuAsn 440
Qy      1348 GACACGAAAGCCGGCTTTTGTCAAGCCGACCAAGCTGATGAAAGTCCGTTGTGTCGT 1407
Db      441 AspThrArgProAlaPheValIySylAlaAspHisSerAspGlnIleIleArgPheValPheGly 460
Qy      1408 GTTGCTTCTCGAAGGGGAGCACTTGTATGTTCCAGAGACCAAGCAGAGAGAGAAAGTTA 1467
Db      461 GlnAlaPheLeuIySylAspIleValIleMetPheGlnIyAlaThrGlnGlnIySylLeu 480
Qy      1468 CTGAGCCGGAAGATGATTAATCTAGGCTACCTTGTCTCGAACCGGGAATCCTAATGG 1527
Db      481 LeuSerArgIyMetMetArgTyIlePheAlaAsnPheAlaArgIleIyAspProAsnIy 500
Qy      1528 AACGACCTGCTCTGTGGCCAGCTTAAATCTGACTGAGACAGTACCTTCAGCTGAGCTTG 1587
Db      501 GlnIyIyValProLeuIyProAlaIyTrThrGlnSerGlnIyIyLeuIyLeuAspLeu 520
Qy      1588 AACATGAGCCCTCGAGACGAGCTCAAGAACCCGGGGTGGATTTTGGACACAGACCATC 1647
Db      521 SerValSerValGlnIySylLeuIySylGlnIyGlnIyValGlnPheIyMetCAsnThrIle 540

RESULT 12
ADF50147
ID      ADF50147 standard; protein, 542 AA.
XX
AC      ADF50147;
XX
DT      12-FEB-2004 (first entry)
XX
DE      Cat cauxin protein SEQ ID NO:4.
XX
KW      cat; cauxin; cat kidney disease marker; kidney disease.
XX
OS      Felis catus.
XX
PN      JP2003250575-A.
XX
PD      09-SEP-2003.
XX
PF      04-MAR-2002; 2002JP-00057908.
XX
PR      04-MAR-2002; 2002JP-00057908.
XX
PA      (TOHO-) TOHOKU TECHNOARCH KK.

```

```

XX      MP1: 2004-002277/01.
DR      N-PSDB; ADF50146.
XX
PT      Novel cauxin protein or its salt, useful as a cat kidney disease marker,
PR      and for diagnosing cat kidney disease.
XX
PS      Disclosure: SEQ ID NO 4; 33pp; Japanese.
XX
CC      The present sequence represents a cat cauxin protein (I) or its salt,
CC      which is cat kidney disease marker. Also described: (1) a partial peptide
CC      (II) of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector
CC      (IV) comprising (III); (4) a transformed host (V) comprising (III) or
CC      (IV); (5) producing (I) or (II) by culturing (V); (6) a antibody (VI)
CC      which couples specifically with (I) or (II); (7) diagnosing cat kidney
CC      disease which involves measuring (I) quantitatively, and where reduction
CC      of amount of (I) indicates presence of the disease; (8) a cat kidney
CC      disease diagnostic agent comprising (I) labelling agent; a reagent which
CC      measures the biological activity of urinary (I) or (VI); and (9) a cauxin
CC      detection kit which measures cauxin in a test sample. (I) is useful as a
CC      cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC      disease. (I) enables detection of cat kidney disease simply and
CC      correctly. (I) provides an early marker for the disease, and replaces
CC      complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC      blood testing.
XX
SQ      Sequence 542 AA:

Alignment Scores:
Pred. No.:      1e-181      Length:      542
Score:          1985.00      Matches:      373
Percent Similarity:      83.27%      Conservative:      60
Best Local Similarity:      71.73%      Mismatches:      85
Query Match:      62.62%      Indels:      2
DB:              8          Gaps:      2

US-10-023-515-3 (1-1746) x ADF50147 (1-542)

Qy      91 GGGCTTCCTGCTAAGGCGCCACAGAGAACACCGAGCTGGATGATTCAGGGCAAGCA 150
Db      22 GlnProAlaIaAspAlaProValArgSerThrArgLeuGlyIlePValArgGlySylGln 41
Qy      151 GTCACTGTGTGGGAAGCCCTGTGCTGAGAGCTGCTCCGAGAGCCCTTGTGCTGCT 210
Db      42 ThrThrValLeuLeuIySerThrIyValProValAsnMetPheLeuGlyIleProIyAlaAla 61
Qy      211 CCCCCGTGGGATCCCTGCGATTACGAACCCGACCTGCGATCCGCTGGAGTAACTTG 270
Db      62 ProProLeuGlyIyProLeuArgPheIySylProIyProAlaLeuProGlnIyAsnAspPhe 81
Qy      271 CGAGAGCCACCCCTCAACCTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTAAAT 330
Db      82 ArgAsnAlaThrSerTyIyProIySylLeuCysPheGlnAspLeuGlnIyIlePValIleSerTyIy 101
Qy      331 CAACACATGCTCAAGGTGATTCACCGAAATTCGAGATGCGAAGAGCTGAGCTCTACCTG 390
Db      102 GlnHisValLeuIySylAlaArgTyIyProIySylLeuGlnIyAlaSerGlnIyAspCysLeuTyIyLeu 121
Qy      391 AACATCTAATGCGCTTCCCAAGCCGATACAGAGCTTCAGACTCCCGCTTGTGGTTC 450
Db      122 AsnIleTyIyAlaProAlaHisIaIaAspAsnGlySerAsnLeuProValIleValIleTyrPhe 141
Qy      451 CCAGAGGTGCTTCAAGACTGGCTCAGCTCCATCTTGTATGGGTCCCGCTGCTGGCC 510
Db      142 ProGlnIyAlaPheIySylMetGlySerAlaSerPheAspGlySerIleAlaValAlaIa 161
Qy      511 TATGAGACGTGCTGCTGTGCTGCTGCTCAGTACCGGCTAGATATATTTGTTCTTACCC 570
Db      162 TyIyGlnAspValLeuIleValIleThrIyGlnIyIyArgLeuGlyIlePheGlyIyPheAsp 181
Qy      571 ACATGGGATCAGACATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGCTGTGCC 630
Db      182 ThrGlnAspGlnHisIaIaArgIyAsnIyIlePValIleLeuIySylGlnIyValAlaIaIleuThr 201

```


CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABB60965-ABB61019 represent novel human
 CC proteins of the invention

XX Sequence 356 AA:

Alignment Scores:

Pred. No.:	4.7e-162	Length:	356
Score:	1780.00	Matches:	340
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	56.15%	Indels:	0
DB:	5	Gaps:	0

US-10-023-515-3 (1-1746) x ABB61006 (1-356)

```

QY 724 ATACTGTCTCCCATGAGCCCAAGGCTTATTCACAAAGCCATGAGAGTGGGGTGGCC 783
Db 17 lLeuSerProMetAlaIysGlyLeuPheH1slyValaIleMetGluSerGlyValaIa 36
QY 784 ATCATCCCTTACCTGAGAGCCCATGATTTATGAGAAAGTGAAGACCTTGAGGTGGTTGCA 843
Db 37 lIleIleProTyLeuGluValaH1sAspTyrgIuYsSerGluAspLeuGlnValaIaIa 56
QY 844 CATTTCTGTGTGAACATGCGTCAGACTCTGAGGCGCTGAGGCTGAGGCTGAGCAAA 903
Db 57 HisPheCySgIyAsnAsnAlaSerAspSerGluAlaLeuLeuAaGySyeuDuYrgH1ys 76
QY 904 CCTTCGAAGAGCTGCTGACCTCAGCCAGCAAAAGATCTTTGACTGAGTGGTTGAT 963
Db 77 ProSerIySgIuLeuLeuThrIeuSerGlnYsThrIySerPheThrArgValaIaAsp 96
QY 964 GATGCTTCTTTCCATGAGCGCTAGATCTATGTTCTGCAAAAGCAATTTAAAGCAATT 1023
Db 97 GlYalAspPhePheProAsnGlnProLeuAspLeuSerGlnYsAlaPheIySAlaIle 116
QY 1024 CCTTCATCATCGAGTCAATTAACACAGTGTGGCTTCTGTCGCTATGAGAGGAGCT 1083
Db 117 ProSerIleIleGlyValaAsnAsnH1sGluCySgIyPheLeuLeuProMetIySgluAa 136
QY 1084 CCTGAGATCTCTGAGTGGCTCCAAAGATCCCTTGCTTCATCTGANTCAAAACATCTCTG 1143
Db 137 ProGluIleLeuSerGlySerAsnIySerIeuAlaLeuH1sLeuIleGlnAsnIleLeu 156
QY 1144 CACATCCCGGCTCAGTATTTGACCTTGAGTAAATGAACTTCCTGATGAGCAAGACATCC 1203
Db 157 HisIleProProGlnIyLeuH1sLeuValaIaAsnGlnIyZrPheH1sAspIyH1sSer 176
QY 1204 CTGACTGAATCCGAGACAGTCTTCTGAGCTTGGAGATGTGTTCTTTGTGTGCTCCT 1263
Db 177 LeuThrGluIleAspSerIeuLeuPheLeuGluYAspValaPhePheValaIaPro 196
QY 1264 GCACGTGATCAAGCTCCATATCACAGAGATGCTGTGACCTGTCTACTTTATGAGTTT 1323
Db 197 AlAleuIleThrAlaArgIyH1sIAspAspAlaGlyAlaProValaIyZrPheTyGluPhe 216
QY 1324 CGGACCGGCGCTCAGTCTTTGAGACACAGAACGCGGCTTTTGCAAGCGACACAGCT 1383
Db 217 ArgH1sArgProGlnCySgPheGluAspThrIySerProAlaPheValaIySAlaAspH1sAla 236
QY 1384 GATGAAGTCGGCTTGTGTGTGCTGTGCTTCTGTAAGGGAGCATTTGTTATGTTGAA 1443
Db 237 AspIuValaIyArgPheValaPheGlyGlyAlaPheLeuYsGlyAspIleValaMetPheGlu 256
QY 1444 GAGGCCAGGAGAGAGAGATTACTGAGCCGGAAGATGATGAATCTGGGCTAAGCTTT 1503
Db 257 GlYAlaIAspThrIuGluGluYsLeuLeuSerIyGlySAspMetCylsTyTrpAlaIAspPhe 276
QY 1504 GCTCGAAGCGGAGATCTTAATGGAGACAGACTGTCTGTGGCCAGCTTAAATCTGACT 1563
Db 277 AlAspArgThrIyAsnProAsnGlyAsnAspLeuSerIeuTrpProAlaIyZrAsnLeuThr 296

```

```

QY 1564 GAGCAGTACTCTCCAGCTTGACTTGAACATGAGCCTCGAGACAGACTCAAGAACCCGCG 1623
Db 297 GluGlnTyLeuGlnIleuAspLeuAsnMetSerLeuGlnIAspGlyGluProArg 316
QY 1624 GTGATTTTGGACACAGACCATCCCCCTGATCCGTCGTGCTCGACATCTCCACAGT 1683
Db 317 ValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuH1sSer 336
QY 1684 CCTCTTTCTTCTTAACTTTCTCTCTCTCTCCAGGCTTTCTTTTCTTTGTGCTCT 1743
Db 337 ProLeuSerSerIeuThrPheLeuSerIeuGlnProPhePhePheAspCyAlaIaPro 356

RESULT 14
AAE20911
ID AAE20911 standard; protein; 561 AA.
AC AAE20911;
XX
XX
DT 01-JUN-2002 (first entry)
XX
XX
DE Rat carboxylesterase-like enzyme protein.
XX
XX
KW Rat; carboxylesterase-like enzyme; organophosphorus intoxication;
KW osteopathic; gene therapy; osteoporosis; antilease therapy; cytostratic;
KW detoxifying agent; Paget's disease; bone implant degradation; cancer;
KW dental implant; enzyme; gene expression.
XX
OS Rattus sp.
XX
XX
PN M0200206454-A2.
XX
PD 24-JAN-2002.
XX
PP 10-JUN-2001, 2001WO-BP007919.
XX
PR 17-JUN-2000; 2000US-0218564P.
XX
XX
PA (FARB ) BAYER AG.
XX
PI Xiao Y;
XX
DR WPI; 2002-195808/25.
XX
PT Novel human carboxylesterase-like enzyme polypeptide, regulators of which
PT are useful for preventing and treating organophosphorus intoxication.
XX
PS cancer and osteoporosis.
XX
PS Disclosure; Fig 8; 92pp; English.
XX
XX
XX The invention relates to a purified human carboxylesterase-like enzyme
XX polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
XX screening for agents which decrease or modulate the activity of
XX carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
XX useful for treating a carboxylesterase-like enzyme dysfunction related
XX disease, such as organophosphorus intoxication, cancer and osteoporosis.
XX Compounds that increase the ability of human carboxylesterase-like enzyme
XX to bind to organophosphorus compounds are useful as detoxifying agents.
XX Carboxylesterase-like enzyme agonists and antagonists are useful for
XX treating osteoporosis. Paget's disease and degradation of bone implants,
XX particularly dental implants. Carboxylesterase-like enzyme is useful in
XX diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to diseases or abnormalities related to the presence of
XX mutations in the nucleic acid sequences which encode the enzyme. The
XX coding sequence of carboxylesterase-like enzyme polynucleotide is useful
XX in gene therapy and for generating antisense oligonucleotides or
XX ribozymes which specifically bind to mRNA transcribed from
XX carboxylesterase-like enzyme DNA. These antisense oligonucleotides are
XX useful for modulating carboxylesterase-like enzyme gene expression. The
XX present sequence is rat carboxylesterase-like enzyme protein
XX
SQ Sequence 561 AA:

```

Alignment Scores:

Pred. No.: 4,886-112 Length: 561
 Score: 1262.50 Matches: 275
 Percent Similarity: 62.41% Conservative: 77
 Best Local Similarity: 48.76% Mismatches: 179
 Query Match: 39.83% Indels: 33
 Gaps: 13

US-10-023-515-3 (1-1746) x AAE20911 (1-561)

```

QY 1 ATGCCACAGGGA---CTTACTTCATCTGCTTCACAAATGGCTTTTCTGATTTCTCAG 57
DB 1 MetProArgaenGlnLeuHisSerTrpLeuSnaIaValLeuPheGlyLeuLeuLeu--- 19
QY 58 CCCGTTGGGACACAGACAGTGGGAAAACCTGGGCTTCCTCGAAGGGCCACAGAG 117
DB 20 ---LeuLeuIleHisValGln---GlyGlnSerProGluSer---SerProIleArg 36
QY 118 AACACACAGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGAGCCCTGTGCT 177
DB 37 ThrThrHisThrGlnGlnValGlnGlyLeuLeuAspHisValLeuAspThrIleValaGly 56
QY 178 GTGAAGTGTTCCTGAGATCCCTTGTCTGCTCCCGCTGGAGATCCCTGAGATTACG 237
DB 57 ValHisThrPheLeuGlyIleProPheAlaLysProProValGlyProLeuArgPheAla 76
QY 238 AACCCGACGCTGCATCCGCTGGGATTAATTGCGAAGAACCACTCTACCTTAATTG 297
DB 77 ProProGluProProGluProTrpSerGlyValArgAspAlaThrSerGlnProAlaMet 96
QY 298 TGCCTCCAGACTCAGTGGCTG-----CTTTAGATCAACAACATGCTCAAG 345
DB 97 CysLeuGlnAsnLeuAspIleLeuAspGlyValGlyLeuLeuAspMetLeuMetIle--- 115
QY 346 GTGCATTAACCCGAATTCGAGGTGCAGAAATGCTCTTACCTTAACATGTATGCGCT 405
DB 116 -----LeuSerSerIleSerMetSerGlnAspCysLeuGlyLeuAsnValIleAlaPro 133
QY 406 GCCCAGCCGATACAGGCTCCAAAGTCCCGCTTGTGTGTGTTCCAGAGAGTCCCTTC 465
DB 134 AlaHisAlaArgGlnGlySerAsnLeuProValMetValTrpIleHisGlyValAlaLeu 153
QY 466 AAGATGCTGAGCTCATCTTTGATGGAGTCCGCGCTGGCTGAGCTTATGAGACCTGCTG 525
DB 154 ValValGlyMetAlaSerMetCysTrpAspGlySerLeuLeuThrValAsnGlnAspLeuVal 173
QY 526 GTTGTGTCTGTCACATACCGGTAGAAATATTGGTTTCTTACCAACATGGAGTACGAT 585
DB 174 ValValThrIleGlnIleGlyArgLeuGlyValLeuGlyPhePheSerThrGlyAspGlnHis 193
QY 586 GCTCCGGGAAATGGGCTTCAAGGACCAAGTGGCTGTCTCTGCTGCTCCGAGAAAGAC 645
DB 194 AlaArgGlyAsnTrpGlyIleLeuAspGlnValAlaAlaLeuArgTrpValGlnGlnAsn 213
QY 646 ATGCAAGTCTTCGGTGGGAGACCCGCTGTGACATCTTTGGAGAGTCCGCGGAGACC 705
DB 214 IleAlaHisPheGlyGlyAsnProAsnArgValThrIlePheGlyGlnSerAlaGlyGly 233
QY 706 ATAAGTGTTCAGTCTTATCTGTCTCCATGAGCAAGGCTTATTCACAAAGCATC 765
DB 234 ThrSerValSerSerHisValIleSerProMetSerGlnGlyLeuPheHisGlyAlaIle 253
QY 766 ATGAGAGTGGGTGGCCATCATCTCTTAACCTGAGAGCCCATGATTATGAGAAAGTGA 825
DB 254 MetGlnSerGlyValAlaLeuLeuProAspLeuIleSer-----GlnThrSerGln 270
QY 826 GACCCGACG---GTGGTTGACATTTCTGTGTAAACATGGCTGACAGCTGAGGCCCTG 882
DB 271 ThrValSerThrThrValAlaLysLeuSerGlyCysGlyAlaMetCaspSerGlnAlaLeu 290
QY 883 CTGAGAGTCTGAGAGCAAAACCTTCCAGAGAGTGTGAGCCCTGACGACCAAGAAACAAAG 942
DB 291 ValArgCysLeuArgAlaLysSerGlyAlaGlnIleLeuValIleAsnLysValPheLys 310
  
```

```

QY 943 TCTTCACTCAGAGTGTGATGATGCTTTCTTTCTTAATGAGCCCTAGATCTATGTCT 1002
DB 311 MetIleProAlaValAlaAspGlyGlnPheLeuProArgHisProLysGlnLeuLeuAla 330
QY 1003 CAGAAAGCATTTAAAGCAATTCCTTCCATCATCTGAGATCAATTAACCAAGATGGCTTC 1062
DB 331 SerGlnAspPheHisProValProSerIleIleGlyValAlaThrAspGlnIleCysCys 350
QY 1063 CTGCTGCTATG-----AAGAGGCTCCCTGAGATCTCTGCT 1098
DB 351 ThrIleProMetValMetGlyThrAlaGlnIleIleLysGln-----LeuSer 366
QY 1099 GAGTTCACACAGTCCCTGCTCCATGCTGATACAAACATCTGACATCCCGCTCAG 1158
DB 367 ArgGlnAsnLeuGlnAlaValLeuLysAspThrAlaAlaGlnMetLeuLeuProGln 386
QY 1159 TATTTGACCTTTGGCTTAATGATTAATTCTTCATGACAAAGCACTCCCTGATGAAATCGA 1218
DB 387 CysGlyAspLeuLeuMetGlyGlyIleCysMetGlyAsnThrAspAspSerGlnThrLeuGln 406
QY 1219 GACAGTCTTTCGAGACTGCTGAGATGTCTTTGTGCTCCCTGACATGATCAGCT 1278
DB 407 IleGlnIleThrGlnMetMetGlyAspPheLeuPheValIleProAlaLeuGlnValAla 426
QY 1279 CGATATCACAGAGATGCTGTGACCTGTCTACTTCTATGATTTGGGACCGGCTCAG 1338
DB 427 HisPheGlnArg---SerHisAlaProValIlePheIleGlyLeuPheGlnHisAlaProSer 445
QY 1339 TGCCTTGAAGACACAGAACCGGCTTTGTCAAAGCCGACCACTGATGAAGTCCGCTTT 1398
DB 446 TyrPheLysAsnValAlaArgProHisValLysAlaAspHisAlaAspGlyValIleProPhe 465
QY 1399 GTTGTGCTGCTGCTCTCTGTAAGGGGACATGTGTATGTTGGAAGACCAAGGAGAG 1458
DB 466 ValPhe---GlySerPhePheSerGlyMetLysLeuAspPhe-----ThrIleGln 481
QY 1459 GAGAAGTACTGAGCGGAGATGATGAATATCTGAGCTTGTCTGCAACCGGAGAT 1518
DB 482 GlnArgLeuLeuSerArgArgMetCysTrpTrpAlaAsnPheAlaArgGlnIleAsn 501
QY 1519 CCTAATGGAAAGCACTGTCTCTGTGGCCAGCTTATTAATCTGACTGACAGTACCTCGAG 1578
DB 502 ProAsnSerGlnGlyLeuProTyrTrpProAlaLeuAspHisAspGlnGlnIleThrLeuGln 521
QY 1579 CTGGACTTGAACATGAGCTCGACAGACATCAAGAAACCGGGGTGATTTTGGACC 1638
DB 522 LeuAspThrHisProAlaValAlaAspArgAlaLeuLysAlaArgArgLeuGlnPheTrpThr 541
QY 1639 AGCACCATCCCC 1650
DB 542 LysThrLeuPro 545
  
```

RESULT 15
 AAB58981
 ID AAB58981 standard; protein; 549 AA.
 XX AAB58981;
 DE 27-MAR-2001 (first entry)
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 689.
 XX Human; breast cancer; cytotoxic; immunosuppressive;
 XX neurotrophic; neuroprotective; antiviral; antiallergic; hepatotropic;
 XX antidiabetic; antiflammatory; antitumor; vulnerary; anticonvulsant;
 XX antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 XX Addison's disease; allergy; autoimmune haemolytic anemia;
 XX autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 XX cardiovascular disorder; wound healing; neurological disease.
 OS Homo sapiens.

This Page Blank (uspto)


```

Db      63 LysProProValGlyProLeuArgPheSerGluProGlnProProGluProTrpIlySerGly 82
Qy      268 TTGGCGAAGAACCCCTCTTACCTTAATTTGGCTCCAGAACTCAGAGTGGCTGCTTTA 327
Db      83 ValArgAspAlaAlaSerTyrProPheMetCysLeuGlnAspLys-----ValLeu 99
Qy      328 GATCAACATGCTC-----AAGGTGCATTACCCCGAATTCGGAGGTGA 372
Db      100 GlyGlnTyrLeuSerAspAlaIleThrAsnArgLysGluValArgLeuGlnIleSer 119
Qy      373 GAAAGACTGCTTACCTGAACATTAATGCGCTCCAGCCGAGTACAGCTCCAAAGCTC 432
Db      120 GlnAspCysLeuTyrLeuAsnValTyrThrProValSerThrGlnGluGlnIlySerLeu 139
Qy      433 CCGGCTTGTTGGTGGTCCAGAGAGTGGCTTCAAGACTGGCTCAGCTCCATCTTTGAT 492
Db      140 ProValPheValTrpIleHisGlyGlyLeuValSerGlyAlaAlaSerSerTyrAsp 159
Qy      493 GGGTCCGCTGCTGCTATGAGAGAGTGTGTGTGTGTGTCGACAGTACCGGCTAGGA 552
Db      160 GlySerAlaLeuAlaAlaPheAspAsnValValValThrIleGlnTyrArgLeuGly 179
Qy      553 ATATTGGTTTCTTACCAACATGAGATGACATGCTCCGGGGAAGTGGCTTCAAGAC 612
Db      180 IlealagIlyTyrPheSerThrGlyAspLysHisAlaArgGlyAsnTrpGlyTyrLeuAsp 199
Qy      613 CAGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
Db      200 GlnValAlaAlaLeuGlnTrpIleGlnGlnAsnIleIleHisPheArgGlyAspProGly 219
Qy      673 TCTGTGACATCTTTTGGAGAGTCCGCGGAGAGCCATAGTGTGTGTGTGTGTGTGTGTGT 732
Db      220 SerValThrIlePheGlyGlySerAlaGlyValSerValSerAlaLeuValLeuSer 239
Qy      733 CCCATGGCCAAAGCTTATTCACAAAGCCATCAGAGAGTGGGCTGGCCATC---ATC 789
Db      240 ProLeuAlaLysGlyLeuPheHisIlySerAlaIleSerGlnSerGlyThrAlaValArgIle 259
Qy      790 CTTTACCTGAGAGCCCATGATTATGAGAAAGTGAAGAGCTGAGGTGGGACATTTTC 849
Db      260 LeuPheTrpGlu-----GlnProGluGlnGlnAlaGlnIleAlaAla 275
Qy      850 TGTGTGTAACAATGCGTCAGACTGTGAGGCTGCTGAGTGGCTGAGGACAAACCTTCC 909
Db      276 AlaGlyCysGluLysSerSerSerAlaAlaLeuValGlnCysLeuArgGluLysTrpGlu 295
Qy      910 AAGAGCTG-----CTGACCTTCAAGCCAGAAACAAAGCTTTTCACTCGAGTGGTAT 963
Db      296 AlaGluMetGluGlnIleThrLeuLysMetProProMetPheIleSerAlaSerLeuAsp 315
Qy      964 GGTGCTTCTTCTTCTTAATGAGCTCTTATGATCTAATTTGTCAAGAAAGCTTTAAACAAT 1023
Db      316 GlyValPhePheProLysSerProArgGlnLeuLeuSerGlyValIleAsnAlaVal 335
Qy      1024 CTTTCATCATCGAGATCAATTAACACAGAGTGGCTTCTGCTGCTCCT--ATGAAGAG 1080
Db      336 ProTyrIleIleGlyValAsnAsnCysGluTrpIleLeuProArgMetMetLys 355
Qy      1081 GCTCTGAGATCTCCAGTGGCTCCAAAGTCCCTTGGCTCCATCTGATCAAAACATC 1140
Db      356 PheProGluPheTrpGluGlyLeuGlnLysAspAlaAlaArgGlnValLeuGlnSerThr 375
Qy      1141 CTGCAATC-----CCGCTCAATATTGACACTTGTGGCTTAATGAATAC 1185
Db      376 LeuAlaLeuSerPheLysGlyAlaProSerAspIleValAspLeuValTyrAsnGluTyr 395
Qy      1186 TTCCATACACAGACTCCCTGACTGAAATCCGAGACAGCTTCTTGAAGCTTGTGAGAT 1245
Db      396 IleGlyValAlaGluValArgAlaGlnValArgAspGlyLeuLeuAspSerIleAlaAsp 415
Qy      1246 GTGTTCTTGTGGTCTGCACTGATCAAGCTCGATATCAAGAGTGGCTGACACT 1305

```

```

Db      416 ProLeuPheValPheSerAlaValGluValAlaArgHisIleArgAspAlaGlyAsnPro 435
Qy      1306 GTTACTTCTATGACTTTTGGGACCGGCTCAGTGGCTTTGAAAGACACAGACCGGCTTTT 1365
Db      436 ValTyrPheTyrGluPheGlnHisIleArgProSerSerAlaIleGlyValValProGluPhe 455
Qy      1366 GTCAAGCCGACACCGCGATGAAGTCCGCTTGTGTGGTGGTGGTGGTGGTGGTGGTGGT 1425
Db      456 ValIlyAlaAspHisAlaAspGluIleAlaPheValPheGlyLysProPheLeuAlaGly 475
Qy      1426 GACATTTGATGTTCGAAGGACCAAGGAGAGAGAGAAATTAATTAAGAGCCGGAATGATG 1485
Db      476 Asn-----AlaThrGluGluGlnAlaLysLeuSerArgThrValMet 489
Qy      1486 AAATTTCTGGCTACCTTTGCTCGAACCGGGAATCTTAATGGGAAAGCATGTCTCTGTGG 1545
Db      490 LysTyrTrpThrAsnPheAlaArgAsnGlyAsnProAsnGlnGlyGlyLeuValHisTrp 509
Qy      1546 CCAAGTTTAAATCTGACTGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1605
Db      510 ProGlnTyrAspMetAspGluArgTyrLeuGlnIleAspLeuThrGlnLysAlaAlaLys 529
Qy      1606 AAGCTCAAAAGAACCGGCGGTGATTTTGTGACACGACCATCCCCCTGATCTGTGCTGCC 1665
Db      530 LysLeuLysGluArgLysMetGluPheTrpMetGlnLeuThrGlnIleMetSerAsp 549
Qy      1666 TCCGACATGCTCCACAGTCTCTT 1689
Db      550 ArgArgArgLysLysIleThrAspLeu 557

RESULT 2
A34329
60K esterase (BC 3.1.1.-) isoform 2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 22-Jun-1990 #sequence_rev10n 22-Jun-1990 #text_change 09-Jul-2004
C:Accession: A34329
R:Ozols, J.
J: Biol. Chem. 264, 12533-12545, 1989
A:Title: Isolation, properties, and the complete amino acid sequence of a second form of
A:Reference number: A34329; MUID:89308686; PMID:2745458
A:Accession: A34329
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-532 <OZO>
A:Cross-references: UNIPROT:P14943
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:32-517/Domain: cholinesterase homology <CHE>
F:201,430/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 3 086-87
Score: 1231.50
Percent Similarity: 62.19%
Best Local Similarity: 47.83%
Query Match: 38.85%
DB: 2 Gaps: 7

US-10-023-515-3 (1-1746) x A34329 (1-532)
Qy      109 CCAACAGAGAACCAAGCTGGAGTGAATTCAGGAGCAAGCACTGCTGAGGAAGC 168
Db      6 ProIleArgAsnThrHisThrGlyGlnValArgGlySerLeuValHisValGluGlyThr 25
Qy      169 CCTGTGCTGTGAACGTGTTCCTCGAGAGTCCCTTTGGTGGTGGTGGTGGTGGTGGTGGT 228
Db      26 AspAlaGlyValHisThrPheLeuGlyIleProPheAlaLysProProLeuGlyProLeu 45
Qy      229 CGATTTAACGAACCCGAGCTGATCGCTCGGATTAATCACTGAGAGACCAAGCACTCTAC 288
Db      46 ArgPheAlaProProGluProAlaGlnAlaTrpSerGlyValArgAspGlyThrSerLeu 65
Qy      289 CTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTATGATCAACATGCTCAAGGTG 348

```



```

Db      ProAlaMetCysLeuGlnAsn-----LeuAlaIleMetAspGlnAspValLeuLeuLeu 83
Qy      CATTAAC-----CCGAATTCGGAGTCTCAAGAAAGCTGCTTACTTGAACATCTATGG 402
Db      HisPheThrProPheSerIleProMetSerGlnAspCysLeuTyIleuAsnIleTyIle 103
Qy      CCGCCCAAGCGCGATACAGAGCTCCGAAGCTCCCGCTTGATGGTCCAGAGAGTGGC 462
Db      ProAlaHisAlaArgGlnGlySerAspLeuProAlaMetValTyrIleHisGlyGly 123
Qy      TTCAAGACTGGCTCAAGCTCCATCTTGAATGGGTCGCGCTGGCTGCTTATGAGAGCTG 522
Db      LeuThiMetGlyMetAlaSerMetTyrAspGlySerAlaLeuAlaIAspGlnAspVal 143
Qy      CTGGTGGTGGTGGTGGTCACTACGAGTACCGGATGAATATTTGGTTTTCACCAAGGATCAG 582
Db      ValValValIleThrIleGlnTyrArgLeuGlyValLeuGlyPhePheSerThrGlyAspGln 163
Qy      CATGCTCCGGGGAATCGGGCCCTTCAAGGACGAGAGTGGCTGTGCTCGGTGGGTCGAAG 642
Db      HisAlaThrGlyAsnHisGlyTyrTyrLeuAspGlnValAlaAlaLeuArgTyrValGlnTyr 183
Qy      AACATCGAGTTCTTGGTGGGAGCCCAAGCTCTGACCATCTTTGGCGAGTCCGCGGA 702
Db      AsnIleAlaHisPheGlyGlyAsnProGlyArgValIleThrPheGlyGlySerAlaGly 203
Qy      GCGATAGTGTCTTGAATCTTACTGTCTGCTCCCAAGGCCAAAGGCTTATTCACAAAGCC 762
Db      GlyThrSerValSerSerHisValLeuSerProMetSerGlnGlyLeuPheHisGlyAla 223
Qy      ATCATGAGAGTGGGGTGGGCAATCCCTTACCTGAGAGGCCCATGATTAATGAAGAAGT 822
Db      IleMetGlySerLeuValAlaLeuLeuProGlyLeuIleThrSerSerSerGlyValVal 243
Qy      GAGGACCTGCAAGCTGGTGGCAATTTCTGTGGTAAACAATGCTCAAGACTTGAAGCCCTG 882
Db      Ser-----ThrValValAlaAsnLeuSerArgCysGlyGlnValAspSerGlnThrLeu 261
Qy      CTGAGAGTCCCTGAGGACAAACCCCTCCAAAGAGCTGTCGACCTCAAGCCCAAGAAACAAAG 942
Db      ValArgCysLeuArgAlaIleSerGlnGlyGlnMetLeuAlaIleThrGlnValIlePheMet 281
Qy      TCTTTCACTGAGTGGTGGTGGTCTTCTTCTTAATGAGCCCTGATGATTAATTTGCT 1002
Db      LeuIleProGlyValValAspGlyValPheLeuProArgHisPheGlnGlyLeuLeuAla 301
Qy      CAGAAAGCATTTAAAGCAATTCCTTCATCGAGAGTCAATACCAAGAGTGGGCTTC 1062
Db      LeuAlaAspPheGlnProValProSerIleIleGlyIleAsnAspAspGlyTyrGlyTyr 321
Qy      CTGCTGCTCT-----ATGAGAGAGCT 1083
Db      IleIleProGlyLeuLeuLeuAlaIleAspProGlnGlnGlyArgAspArgGlnAlaMet 341
Qy      CCTGAGATCTCAAGTGGCTCCAAAGTCCCTTGCCCTGCACTGATACAAACATCTCG 1143
Db      ArgGlnIleMetHisGlnAlaThrTyrGlnLeuMetLeu----- 354
Qy      CACATCCGCGCTCAGATATTGACCTTGAGCTTAATGATATTCCTCATGACAAAGACTCC 1203
Db      -----ProProAlaLeuGlnValAspLeuMetAspGlyTyrMetGlySerAsnGlyAsp 372
Qy      CTGACTGAATTCGGAAGAGCTTCTGCACTTGTGAGATGTGTCTTTGTGGCTCT 1263
Db      ProGlyHisIleMetAlaGlnPheGlnGlnMetMetAlaAspAlaMetPheValIleMetPro 392
Qy      GCACTGATCACAAGCTCGATATACAGAGATGCTGGTGCACCTGTCTACTCTATGATGTT 1323
Db      AlaLeuArgValAlaHisIleGlnTyrG--SerHisAlaProThrTyrPheTyrGlyPhe 411
Qy      CGGACCGGCGCTCAGTGTCTTGAAGACGAGAGCCGCGCTTTGTCAAGCCGACGACGCT 1383

```

```

Db      GlnHisArgProSerPheThrTyrAspLeuArgProProHisValArgAlaAspHisGly 431
Qy      GATGAAGTCCGCTTGTGCTGGTGGTGGCTCTTCTGAAGAGGGGACATTGTTATGTCGA 1443
Db      AspGlnValValPheValPheArgSerHisValLeuPheGlySerLeuValProLeu----- 449
Qy      GAGGACGAGGAGGAGAGAAATTAATGAGCCGGAAGATGATGAATTAATGAGCTTAT 1503
Db      -----ThrGlnGlnGlnGlnLeuLeuSerArgValMetLeuTyrTyrIleAsnPhe 467
Qy      GTCGACCCGCGGATTCCTTAATGGGAACAGCTGTCTGTGGCCGACTTAATTAATGACT 1563
Db      AlaArgAsnArgAsnProAsnGlyGlyGlyLeuAlaHisTyrProLeuPheAspLeuSer 487
Qy      GAGCAGTACCTTCAGTGGAGCTTGAACATGAGCTTCGGACAGAGACTCAAGACGCGG 1623
Db      GlnArgTyrLeuGlnLeuAsnMetGlnProAlaValGlyGlnAlaLeuValAlaArgTyr 507
Qy      GTGATTTTGGACGACGACCATCCCT 1650
Db      LeuGlnPheThrThrHisThrLeuPro 516

RESULT 3
JC5408
carboxylesterase (EC 3.1.1.1) - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: JC5408
R:Schwartz, H.; Langmann, T.; Daig, R.; Becker, A.; Aslanidis, C.; Schmidt, G.
Biochem. Biophys. Res. Commun. 233, 117-120, 1997
A>Title: Molecular cloning and characterization of a novel putative carboxylesterase, pre
A:Reference number: JC5408; MUID:97289502; PMID:9144407
A:Accession: JC5408
A:Molecule type: mRNA
A:Residues: 1-559 <SCH>
A:Cross-references: UNIPROT:000748; GB:Y09616; MID:92058317; PIDN:CAA70831.1; PID:9205831
A:Experimental source: intestine
C:Comment: This enzyme hydrolyzes many xenobiotics, such as carboxyl esters, thioesters &
C:Genetics:
A:Gene: GDB:CEB2; ICE; CE2
A:Cross-references: GDB:9959011
A:Superfamily: Cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:58-544/Domain: cholinesterase homology <CHS>
F:15-95,123-280,291-428/Disulfide bonds: #status predicted
F:111,276/Binding site: carbonylate (Asn) (covalent) #status predicted
F:128,457/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 4,07e-87 Length: 559
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 38.80% Indels: 24
DB: 1 Gaps: 8

US-10-023-515-3 (1-1746) x JC5408 (1-559)
Qy      GGGCGCTTGTCTGAAGGGCCACAGAGAACACGAGCTGGAGTGAATTCAGGGCAAGCA 150
Db      GlnGlnAspSerPheAlaSerProIleArgThrThrHisThrGlnValLeuGlySerLeu 45
Qy      GTTCATGTGTGGGAAAGCCCTGTGCTGGAACGTTCTCTGGAGTCCGCTTGTGCTGT 210
Db      ValHisValIleGlyAlaAsnAlaGlyValGlnThrPheLeuGlyIleProPheAlaVal 65
Qy      CCGCGCTGGATCCCTCGATTTTACGAACCGGACGCTGCACTCCGCTGGGATTAATTTG 270
Db      ProProLeuGlyProLeuArgPheAlaProProGlyProProGlySerThrPheSerGlyVal 85
Qy      CGAGAAGCCACCTCTACCTTAATTTGTGCTCCGAAC-----TCAGAG 315
Db      ArgAspGlyThrThrHisPheProAlaMetCysLeuGlnAspLeuThrAlaValGlnSerGln 105

```

```

QY 316 TGGCTGCTTTAGATCAACATGCTCAAGTGCATTAACCCGAATTCGGAGTGCAGA 375
   |||
   |||
Db 106 PheLeu-----SerGlnPheAsnMetThrPheProSerXpSerMetSerGln 121
QY 376 GACTGCTCTACTGTAACATCTTATGGCGCTGCCCAAGCCGATACAGGCTTCAAGCTCCC 435
   |||
   |||
Db 122 AspCyALeuIlyrLeuSerXpIleTyTrnProAlaHiserXpIleGluIlySerAsnLeuPro 141
QY 436 GTCTTGATGATGTTCCAGAGAGTGCCTTCAAGTGCCTGACGCTCCATCTTTGATGGG 495
   |||
   |||
Db 142 ValMetValTrpIleHisGlyAlaLeuValPheGlyMetAlaSerLeuIlyrAspGly 161
QY 496 TCCGCGCTGGCTGCTATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
   |||
   |||
Db 162 SerMetLeuAlaAlaLeuGluAsnValValValIleIleGlnIlyrTrnGluGlyVal 181
QY 556 TTTGTTGTTTTCACCAACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACGAG 615
   |||
   |||
Db 182 LeuGlyPhePheSerXpThrGlyAspIlyrValAlaThrGlyAsnTrpGlyTrnLeuAspGln 201
QY 616 GTGGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
   |||
   |||
Db 202 ValAlaAlaLeuAlaGlyTrpValGlnGlnAsnIleAlaHisPheGlyGlyAsnProAspArg 221
QY 676 GTGACCATCTTTGGCGAGTCCGCGGAGCCATAGTGTTCATGCTTTACTGCTGCTCCC 735
   |||
   |||
Db 222 ValTrnIlePheGlyGlySerAlaGlyIlyTrnSerValSerSerLeuValValSerPro 241
QY 736 ATGGCCAAAGGTTTATTCACAAAGCCATCAGAGAGAGTGGGCTGCCATCATCCCTTAC 795
   |||
   |||
Db 242 IleserGlnGlyLeuPheHisGlyAlaIleMetGlnSerGlyValAlaLeuLeuProGly 261
QY 796 CTGAGAGCCCATGATTAATGAGAAGAGTGAAGACTG-----CAGGTGGTTCACATTTTC 849
   |||
   |||
Db 262 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAsnLeu 277
QY 850 TGTGTTAAACATGCTCAAGACTCTGAGCGCTGCTGAGGTGCTGAGAGCAAAACCTTCC 909
   |||
   |||
Db 278 SerAlaCyAAspGlnValAspSerGlnAlaLeuValGlyCyALeuAlaArgGlyLysSerLys 297
QY 910 AAGAGCTGCTGAGCTGAGCTGAGCCAGAAACAAAGTCTTTCACGCTGCTGATGTGCT 969
   |||
   |||
Db 298 GlnGlnIleLeuAlaIleAsnIlyrPheLysMetIleProGlyValValAspIlyVal 317
QY 970 TTTCTTTCTTAAGACTCTGATCTATGTTGTTCTCAGAAAGATTAAGCAATTCCTTCC 1029
   |||
   |||
Db 318 PheLeuProArgHisPheProGlnGluLeuLeuAlaSerAlaAspPheGlnProValProSer 337
QY 1030 ATCATCGAGTCAATACCAAGAGTGTGGCTTCTGCTGCT-----ATGAAG 1077
   |||
   |||
Db 338 IleValGlyValAsnAsnAsnGlnPheGlyTrnLeuIleProLysValMetArgIleTyx 357
QY 1078 GAGGCTGCTGAGTCAAGTGCCTGAGGCTCCAAAGTCCCTTGCCTCATCATCAAAAC 1137
   |||
   |||
Db 358 AspTrnGlnIlyrGluMetAspArgGluAlaSerGlnAlaAlaLeuGlnIlyrMetLeuThr 377
QY 1138 ATCTGCAACATCCGCGCTCAGTATTGTCACCTGTGGCTTAATGTAATCTTCATGACAAG 1197
   |||
   |||
Db 378 LeuLeuMetLeuProProThrPheGlyAspLeuLeuArgGluIlyrIleGlyAspAsn 397
QY 1198 CACTCCCTGACTGAATCCGAGACAGTCTTGCATCTTGCATCTTGCATGATGTTGTTGTG 1257
   |||
   |||
Db 398 GlyAspProGlnTrnIleGlnAlaGlnPheGlnIlyrMetMetAlaAspSerMetPheVal 417
QY 1258 GTCCCTGACTGATCAGACGCTGATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
   |||
   |||
Db 418 IleProAlaLeuGlnIlyrAlaAlaHisPhe---GlnCySerHisArgAlaProValIlyrPheTyx 436
QY 1318 GAGTTTGGGACCGCGCTCAGTGTCTTGAAGACAGAGCGCGCTTTTGTCAAGCGCAC 1377
   |||
   |||
Db 437 GlnPheGlnHisGlnProSerTrpLeuLysAsnIleArgProProHisMetIlyrAlaAsp 456

```

```

QY 1378 CAGCGTGAATGATCGGCTTTGTGTGCTGCTGCTTCTGAAAGGAGCATTTGTATG 1437
   |||
   |||
Db 457 HisGlyAspGlnLeuProPheValPhe---ArgSerPhePheGlyGlyAsnTrpIleLys 475
QY 1438 TTGAAAGGAGCCAGAGAGAGAGTAACTGACGCCGGAAGATATGATAATATGAGGCT 1497
   |||
   |||
Db 476 Phe-----ThrGlnGlnGlnIlyrGlnLeuSerArgLysMetLeuIlyrTrpAla 492
QY 1498 ACCTTGTCTCCAAACCGGGAATCTTAATGGAAAGCATGCTGTCTGTGGCCAGCTTAAAT 1557
   |||
   |||
Db 493 AsnPheAlaArgAsnGlyAsnProAsnGlyGlnIlyrLeuProHisTrpProLeuPheAsp 512
QY 1558 CTGACTGAGCAGTACTTCAGCTGAGTGAACATGAGCTCCGAGACAGACTCAAAADA 1617
   |||
   |||
Db 513 GlnGlnGlnGlnIlyrLeuGlnLeuAsnMetGlnProAlaValGlyArgAlaLeuValAla 532
QY 1618 CCGCGGTGATTTTGGACGACGACCATCCCC 1650
   |||
   |||
Db 533 HisArgLeuGlnPheTrpIlyrValAlaLeuPro 543

```

RESULT 4

S47655 carboxylesterase (EC 3.1.1.1) precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S47655
R:Some, T.: Isoe, M.; Takabatake, E.; Wang, C.Y.
Biochim. Biophys. Acta 1207, 138-142, 1994
Article: Cloning and sequence analysis of a hamster liver cDNA encoding a novel putative
A:Reference number: S47655; MUID:94318665; PMID:8043605
A:Accession: S47655
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <SON>
A:Cross-references: UNIPROT:Q64419; EMBL:ID28566; NID:9531238; PTDN:BAA05913.1; PTD:95312;
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:58-546/Domain: cholinesterase homology <CHR>
F:1227,459/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Query Match:	Indels:	Gaps:
1	561	1,036-84	252	74	47.28%	181	37.82%	26	7

US-10-023-515-3 (1-1746) x S47655 (1-561)

```

QY 91 GGGCTTGTGCTGAAGGCGACAGAGACAGAGCTGGATGATTCAGGGCAAGCA 150
   |||
   |||
Db 26 GlyGlnAspSerValSerProIleArgValHisThrIlyrGlnValArgGlyLysLeu 45
QY 151 GTCACTGTGCTGGAAGCCCTGTGCTGATGAGTGTCTTCCGAGATCCCTTGTGCT 210
   |||
   |||
Db 46 ValIlyrValLysGlnGlyValThrGlyValIlyrAlaPheLeuGlyIleProPheAlaLys 65
QY 211 CCCCCGCTGGAATCCCTGCGATTTACGAACCCGAGCTGCATGCGCTGGAGTAACTTG 270
   |||
   |||
Db 66 ProProValGlyProLeuArgPheAlaProGlnIlyrProGlnIlyrProTrpSerGlyVal 85
QY 271 CGAAGACCACTCCCTCAACCTTAATTTGTGCTCCAGACTCAGAGTGGCTGCTTAAGT 330
   |||
   |||
Db 86 ArgAspGlyThrSerGlnProAlaMetCyALeuGlnThrAspPheMetArgProGlnIle 105
QY 331 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAGACTGCTTACCTG 390
   |||
   |||
Db 106 SerIlyrGlnArgValIleIleuProTrnIleSerMetSerGlnAspCyALeuIlyrLys 125
QY 391 AACATCTAAGCGCTGCCACGCGCATACAGGCTCCAGCTCCCGCTTGTGTGCTTC 450
   |||
   |||
Db 126 AsnIleTyTrnProAlaHisAlaHisGlnIlyrSerAsnLeuProValMetValTrpIle 145

```

```

QY 451 CCAGAGGTCCTTCAAGAGTCCTCAGCTCCATCTTGGATGGGCGCCCTGGCTGC 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 146 HtGtGtGtAlaLeuValMeGtGtMeAlaSerMetAlaSerMetGtGtGtAla 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 511 TATGAGAGCTGCTGGTGTGTGCTGCTCAGTACCGGCTAGAGAAATTTGGTTCTTCA 570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 ThrGtGtAlaLeuValIleValSerIleGtGtGtAlaLeuGtGtGtGtGtSer 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 571 ACATGGGATCAGAGCTCTCCGGGAACTGGGCTTCGAAAGACAGGTGGCTCTGTCC 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 ThrGtGtAlaLeuValIleValSerIleGtGtGtAlaLeuGtGtGtGtGtSer 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 631 TGGGTCGAGAGACATGAGTCTTCCGTTGGGAGCCCGACGCTGTGACATCTTGGC 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 206 TrpAlaGtGtAlaLeuValIleValSerIleGtGtGtAlaLeuGtGtGtGtGtSer 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 691 GAGTCCGCGGAGCCATGATGTTTCTAGTCTTATATCTGTCTCCATGGCCAAAGCTTA 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 ValSerAlaGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 751 TTCCCAAGGATATGAGAGAGTGGGCTGCTCCTTACCTTGGAGGCCATGAT 810
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 PheHtGtGtAlaLeuValIleValSerIleGtGtGtAlaLeuValIleValSer 264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 811 TATGAGAGATGAGAGCTGAGGCTGAGTGTGATGCTTCTTCTTATGAGCTCTTA 870
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 ThrProGtAlaValAlaValThrProValAlaAlaValGtGtGtGtGtGtGtGt 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 871 TCTGAGGCTCTGAGGCTGAGGCTGAGCAAAACCTCCAAAGAGCTGAGCCCTCAGC 930
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 285 SerGtGtAlaLeuValIleValSerIleGtGtGtAlaLeuValIleValSer 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 931 CAGAAAACAAAGCTTTCACTGAGTGTGATGCTTCTTCTTATGAGCTCTTA 990
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 305 GlnValPheIleMetThrProGtAlaValAlaValGtGtGtGtGtGtGtGt 324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 991 GATCATATGCTCAGAAAGCAATTTAAAGCAATCTCTTCATGAGAGTCAATACCA 1050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 325 GlnLeuLeuAlaSerValAlaPheIleProValProSerIleIleGtGtAla 344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1051 GAGTGTGCTCTCTGCTGCTATG-----AAGAGAGCT 1083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 345 GlnGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1084 CTTGAGATCTTCACTGCTCCAAAGTCTCTGCTCCATCTGATACAAACATCTCTG 1143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 365 ArgGlnThrLeuProAlaPheLeuIleSerArgAlaGlnIleMetMet----- 380
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1144 CACATCCCGCTCAGATTTGCACTTGTGCTAATGAAATCTTCCATGAGCAAGACTCC 1203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 ---LeuProProGlnGtGtSerAlaPheLeuMetGlnGtGtGtAlaPheAla 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1204 CTGACTGAATCTCCGAGAGTCTTCTGAGATGCTTGGAGATGTTCTTGTGGCTCT 1263
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 400 ProGlnThrLeuGlnAlaGlnPheArgGtGtGtGtGtGtGtGtGtGtGtGtGt 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1264 GCATGATCAGAGCTGATGATCAGAGATGCTGTGCACTGTCTTCTTATGAGTTT 1323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 AlaLeuValAlaValAlaValPheGlnArg---SerHtAlaProValAlaVal 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1324 CCGGACCGGCTCAGTCTT-----GAGACACGAAAGCGCTTTGTCAACCGAC 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 GlnHtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1378 CAGCTGATGAGTCCGCTTGTGCTGCTGCTCTC-----CTGAAGGGAGCAAT 1431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 HtGtGtAlaPheAlaValAlaPheGtGtGtGtGtGtGtGtGtGtGtGtGt 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1432 GTTATGTTGAAAGGAGCGAGAGAGAAATTAAGTACGCGAGAGATGATGAATAC 1491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 -----ThrGlnGlnGlnGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 1492 TGGGCTACCTTTGCTGGAACCGGAGATCTTAATGGAAGACAGCTGTCTGTGGCAGCT 1551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 493 TrpAlaAlaPheAlaAlaArgHtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1552 TATAATCTGACTGAGCAGTACTCCAGTGTGAGCTTGAACATGAGCTCGAGACAGACTC 1611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 513 LeuValHtAlaPheArgGlnThrLeuValSerIleGlnProAlaValAlaVal 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1612 AAAGAACCGCGGATGATTTTGGACACAGCAACATCCCC 1650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 LysSerArgLysLeuHtPheThrThrLysIleLeuPro 545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
S34607
carboxylesterase (EC 3.1.1.1) - mouse
C.Species: Mus musculus (house mouse)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C.Accession: S34607
R.Aida, K.; Moore, R.; Negishi, M.
Biochim. Biophys. Acta 1174, 72-74, 1993
A.Title: Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase in
A.Reference number: S34607; MUID:93326638; PMID:7916639
A.Accession: S34607
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-554 <AID>
A.Cross-references: UNIPROT:O63880; GB:S64130; NID:g404388; PIDN:AA827606.1; PID:g404389
C.Superfamily: cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase
F.46-536/Domain: cholinesterase homology <CH>
F.215.443/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 9.57e-84 Length: 554
Score: 1186.50 Matches: 256
Percent Similarity: 61.16% Conservative: 81
Best Local Similarity: 46.46% Mismatches: 183
Query Match: 37.43% Indels: 31
DB: 1 Gaps: 10

US-10-023-515-3 (1-1746) x S34607 (1-554)
QY 37 TGCCTTTTCGATATCTCAGCCCTGTGGACACAGAGTGGGAAAAAGCTGGCGCT 96
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 CysLeuLeuLeuIlePheProThrThrVal-----IleGtPro 15
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 97 TCTGTCGAGGCGCACAGAGAACACAGAGCTGGAGTATTCAGGCGAAGCAACTCACT 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16 LysValThrGlnProGlnValAlaPheThrProLeuGlnValAlaArgGlnValGly 35
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 157 GTGCTGGAGAGCCCTGTGCTGTGAAGCTGTTCTCGGAGTCCCTTGTGCTGCTCCCGC 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36 ValLysAlaPheThrAlaPheArgMetValAlaValPheLeuGlnIleProPheAlaGlnAlaPro 55
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 217 CTGGGATCCCTCGGATTTACGAAACCGCAGCTGATGAGCCCTGGAGTAATCTGGAGAA 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56 LeuGlyProLeuAlaPheSerAlaProLeuProProGlnProGlnProGlnGlyAlaArg 75
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 GCCACTCTCACTCAATTTGTGCTCCAGAACTGAGAGTGGCTGCTTATGATCAACAC 336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 AlaSerIleAlaProProMetCysLeuGlnAlaValGlnArgMetSerAlaSerArgPhe 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 ATGCTCAAGGTGATTAACCCGAAATTTGGAAGTGTGAAGAGTGCCTTACTGAAACATC 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 ThrLeuAlaGlnLysMetLysIlePheProIleSerGlnAlaPheCysLeuThrLeuAlaIle 115
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 397 TATGCGCTGCGCCACGCGGATACAGGCTCCCAAGCTCCCGTGTGATGATGTTCCAGGA 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 116 TyrSerProThrGlnIleThrAlaGlnAlaPheAlaPheValAlaThrIleHtGly 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 457 GGTGCTTCAAGAGTGTGCTCAGCTCTCATCTTTGATGGGTCGCGCTGAGCTGCTATGAG 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 GlySerLeuArgValGlySerSerThrSerHtAlaPheGlySerAlaLeuAlaIleArgGly 555
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 517 GACGTGCTGTTGGTCCGACAGTACCGGCTAGAAATATTTGGTTCTTCCACACATGG 576
Db 156 AapValValValValThrValGlnTyrAlaGlnGlyIlePheGlyLeuLeuSerThrGly 175
QY 577 GATCAGCATGCTCCGGGAACTGGGGCTTCAGAGACAGAGTGGCTGTGTCTGGAGTC 636
Db 176 AapLysHleMetProGlyAenArgGlyPheLeuAapValValAlaHleuHArgTrpVal 195
QY 637 CAGAAAGACATGAGTTCTTGGTGGGACCCGACGTCGTGATCACTTTGGCCGAGTCC 696
Db 196 GlnGlyAenHleAlaProPheGlyIleAapProAenCysValThrIlePheGlyAenSer 215
QY 697 GCGGAGACCATTAAGTTTCTAGCTTATATCTGTCTCCATGGCCCAAGGCTTATTTCCAC 756
Db 216 AlaGlyGlyIleGlyIleValSerSerLeuLeuSerProMetSerAlaGlyLeuPheHis 235
QY 757 AAGGCTATCATGAGAGATGGGGTGGCCATCATCCCTTACCTGAGGCGCATATATATGAG 816
Db 236 ArgAlaHleSerGlnSerGlyValValIleSerLysIleLeuGlu-----AapLeuAen 253
QY 817 AAGAGTGAAGACCTGACGAGTGGTGGACAT-----TTCTGTGTAAACATGCGTCAGAC 870
Db 254 AlaTrpSerGlnAlaGlnAenPheAlaAenSerValAlaCysGly---SerAlaSerPro 272
QY 871 TCTGAGGCGCTGCTGAGGTGCTGAGACAAACCTCCAGAGACCTGTGACCTTCAGC 930
Db 273 AlaGlu---LeuValGlnCysLeuLeuGlnLysGlnGlyLysAapLeuHleThrLysLys 291
QY 931 CAGAAACAAAGTCTTACCTGAGATGGTGTGATGTGCTTTCTTCTTATAGCTCTCA 990
Db 292 AenValAenHleSerLysTrp-----ValAenAapSerPhePheProGlnArgProGln 309
QY 991 GATCATTTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACAC 1050
Db 310 LysLeuLeuAlaAenLysGlnPheProThrValProTyrLeuLeuGlyValThrAenHis 329
QY 1051 GAGTGTGCTTCTGCTG-----CCTATGAGAGG---GCT 1083
Db 330 GlnPheGlyTrpLeuLeuLysPheTrpAenHleLeuAapLysMetGlnHleLeuSer 349
QY 1084 CCGTGAATCTTCAGTGGGCTCCAAAGACCCCTGGCCCTTCATGTATACAAACATCTGTG 1143
Db 350 GlnGlnAapLeuLeuGlnAenSerArgProLeuLeuAlaHis-----Met 364
QY 1144 CACATCCGCGCTCAGTATTTGACCTTGGCTATGATGAATATCTTCATGACAAAGCATCTCC 1203
Db 365 GlnLeuProProGlnHleMetProThrValIleAapGlyTyrLeuAapAenGlySerAap 384
QY 1204 CTGACTGAATCCGAGACAGTCTTGTGACCTTGTGAGATGTGTTCTTTGGTCCCTT 1263
Db 385 GlnSerAlaThrArgTyrAlaLeuGlnGlnLeuLeuGlyAapIleThrLeuValIlePro 404
QY 1264 GCACGTATCAGACGCTCGATATCACAGAGATGCTGTGACCGCTGTCTTCATGAGAGTT 1323
Db 405 ThrLeuHlePheSerLysTyrLeuGlnAapAlaGlyCysProValPheLeuTyrGlnPhe 424
QY 1324 CGGACCGCGCTCAGTCTTTGAAGACAGAGCGGCTTTGTCAAGCCGACACACGCT 1383
Db 425 GlnHleThrProSerSerPheAlaLysPheLysProAlaTrpValLysAlaAapHisSer 444
QY 1384 GATGAAGTCCGCTTTGGTTCGGTGGTCCCTTCTG-----AAGGGGACATTTGTTATG 1437
Db 445 SerLysAenAlaPheValPheGlyGlyProPheLeuThrAapGlnSerSerLeuLeuAla 464
QY 1438 TTCCAGAGGACGACGAGAGAGAGAAATTACTGACCGGAGAGATTAATACCGGCTT 1497
Db 465 PheProGlnAlaThrGlnGlnGlnGlnLysGlnLeuSerLeuThrMetMetAlaGlnTrpSer 484
QY 1498 ACCTTGTCTGACACCGGAGATCTTAATGGAACACCTGTCTGTGGCCAGCTTAAT 1557
Db 485 GlnPheAlaArgTrpGlnLysAenProAenGlyLysGlyLeuProTyrProGlnLeuAen 504

```

```

QY 1558 CTGACTGACAGTACTTCGACGCTTGAAATGAGCTTCGACAGAGCTCAAGAA 1617
Db 505 GlnLeuGlnGlnTyrLeuGlnGlnIleGlyLeuGlnLysProAArgThrGlyValLysLeuLysLys 524
QY 1618 CCGGGGTGATTTTGGACGACGACCATCCCC 1650
Db 525 GlyArgLeuGlnPheTrpThrGlnThrLeuPro 535

RESULT 6
A39060
carboxylesterase (EC 3.1.1.1) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 09-Jul-2004
C:Accession: A39060
R:Ovnic, M.; Tepperman, K.; Medda, S.; Elliott, R.W.; Stephenson, D.A.; Grant, S.G.; Gane
Genomics 9, 344-354, 1991
A:Title: Characterization of a murine cDNA encoding a member of the carboxylesterase mult
A:Reference number: A39060; MUID:91169540; PMID:1840565
A:Accession: A39060
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <OWN>
A:Cross-References: UNIPROT:P23953; GB:M57960; NID:g192853; PID:AAA6397.1; PID:g192854
C:Superfamily: Cholinesterase; Cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:50-540/Domain: cholinesterase homology <CHB>
F:221,455/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 6,35e-82 Length: 554
Score: 1163.00 Matches: 241
Percent Similarity: 59.74% Conservative: 84
Best Local Similarity: 44.30% Mismatches: 193
Query Match: 36.69% Indels: 26
DB: 2 Gaps: 7

US-10-023-515-3 (1-1746) x A39060 (1-554)
QY 46 CTGATCTCCAGCCCGCTTGGAGACAGACAGTGGGAAAACTGGGCGCTTGTGTGA 105
Db 11 LeuAlaValCysProIleLeuGlyHis-----SerLeuLeu 22
QY 106 GGGCCACAGAGAAACCCAGGCTGGATGATTCAGGCGAAGCAAGTCACTGTGCGGA 165
Db 23 ProProValValAapThrTrpGlnGlnGlyValValLeuGlnLysTyrIleSerLeuGlnGly 42
QY 166 AGCCCTGCTGCTGGAACGCTTCCCTGGAGTCCCTTGTGCTGCTCCCGCTGGATCC 225
Db 43 PheGlnGlnProValAlaValAlaPheLeuGlyValProPheAlaLysProProLeuGlySer 62
QY 226 CTGGCATTTAAGAACCCGACGCTGCATCCCGCTGGAGTAACTTGCGAGAGAACCATCTCC 285
Db 63 LeuArgPheAlaProProGlnProAlaGlnProIlePheValLysAenAlaThrSer 82
QY 286 TACCTTAATTTGGCTCCGAACTCAGAGTGG-----CTGCTTTAATCAACACATG 339
Db 83 TyrProMetCysSerGlnAapAlaGlyTyrAlaLysIleLeuSerAapMetPheSer 102
QY 340 CTCAAGGTGATTAACCGGAATTCGAGTGTGAGAGACGCTTCATCGAGACATATAT 399
Db 103 ThrGlnLysGlnLeuLeuProLeuLysIleSerGlnAapCysLeuTyrLeuAenIleTyr 122
QY 400 GCGGCTGCGCACGCGCATACAGGCTCCAGCTCCCGCTTGTGGTGTGTTCCGAGAGGT 459
Db 123 SerProAlaAapLeuThrLysSerSerGlnLeuProValMetValTrpIleHleGlyGly 142
QY 460 GCCTCAAGCTGGCTCAGGCTCCATCTTTGATGGGTCCCGCTGGCTGCTATAGAGAC 519
Db 143 GlyLeuValIleGlyArgSerProTyrAenGlyLeuAlaLeuSerAlaHleGlnAen 162
QY 520 GTGCTGTTGTGGTCCGAGTACCGGATGAGAAATTTGGTCTTCTTCACCAATGGGAT 579
Db 163 ValValValValThrIleGlnTyrAlaGlnGlyIleTrpGlyLeuPheSerThrGlyAap 182

```

```

QY 580 CAGATGCTCCGGGAACTGGCTTCAAGACGAGTGGCTGTCTGTCCTGGGTCG 639
DB 183 GlnHisSerProGlyAsnThrPalaHisLeuAspGlnLeuAlaLeuArgTrpValGln 202
QY 640 AAGAACATGAGTTCTTCGTGGGAGCCCACTGTGACCATCTTGGCGAGTCGGG 699
DB 203 AsnAsnIleAlaAsnPheGlyGlyLeuAsnProAspSerValThrIlePheGlyGlySer 222
QY 700 GGAGCCATTAAGTTCTTAATCTTAATCTGTCTCCCAAGCCAAAGGCTTATTCACAA 759
DB 223 GlyGlyIleSerValSerValLeuValLeuSerProLeuGlyLysAspLeuPheHisArg 242
QY 760 GGCATCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 819
DB 243 AlaIleSerGlnSerGlyValValIleAsnThrAsnValGlyLysAsnIleGlnAla 262
QY 820 AGTGAGACCTGACAGTGGTGGACATTTCTGTGGTAAACATGCGTCAAGCTCGAGGCC 879
DB 263 ValAsnGlnIleIleAlaThrLeuSerGlnCys-----AsnAspThrSerSerAlaIa 280
QY 880 CTGCTGAGAGTGCCTGAGACAAACCTCCAAAGAGCTGTCGACCTCAGC----- 930
DB 281 MetValGlnCysLeuArgGlnLysThrGlnSerGlnLeuLeuGlnIleSerGlyLysLeu 300
QY 931 ---CAGAAACAAAGCTTTCCTGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 987
DB 301 ValGlnTyraAsnIleSerLeuSerThrMetIleAspGlyValValLeuProLysAlaPro 320
QY 988 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATAC 1047
DB 321 GlnGlnIleLeuAlaGlnLysSerPheAsnThrValProTyrlleValGlyPheAsnLys 340
QY 1048 CACGAGTGTGCTTCCTGCTGCTTATG-----AAGAGAGCTCCTGAG--ATCCTC 1095
DB 341 GlnGlnPheGlyTrpIleIleProMetMetLeuGlnAsnLeuLeuProGlnLysMet 360
QY 1096 AGTGCTCCAAACAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1155
DB 361 AsnGlnGlnThrAlaSerLeuLeuLeuArgArgPheHisSerGlnLeuAsnIleSerGln 380
QY 1156 CAGTATTGACCTTGTGGCTAATGATATTCTTCATGACAGACACTCCCTGATGTAATC 1215
DB 381 SerMetIleProAlaValIleGlnGlnTyrlleuArgGlyValAspAspProAlaLysLys 400
QY 1216 CGAGACAGTCTTCTGACTGCTTGGAGATGTGTTCTTGTGTCCTGCTGCTGCTGCTG 1275
DB 401 SerGlnLeuIleLeuAspMetPheGlyAspIlePhePheGlyIleProAlaValLeuLeu 420
QY 1276 GCTGATATCACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335
DB 421 SerArgSerLeuArgAspAlaGlyValSerThrTyrlleuArgGlyPheArgTyrlleuPro 440
QY 1336 CAGTCTTGAAGACAGACAGCGCGCTTTTGTCAAGCGGACAGACAGCTGATGAATCGC 1395
DB 441 SerPheValSerAspLysArgProGlnThrValGlnGlyAspHisGlyAspGlnIlePhe 460
QY 1396 TTTGTGTTGGTGGTGGCTTCTGCTGAGAGGAGCATTTGTTATGTTGAAGAGCCAGAG 1455
DB 461 PheValPheGlyAlaProLeuLeuLys-----GlnGlyLysSerGln 474
QY 1456 GAGGAGAAAGTTACTGACCGGAGAAAGATGATGAATACTGGGCTACTCTTGTCTGAAACGG 1515
DB 475 GlnGlnThrAsnLeuSerLysMetValMetLysPheTrpAlaAsnPheIleArgAsnGly 494
QY 1516 AATCTTAATGGAAGACACTGTCTCTGAGCGAGCTTATTAATCTGATGAGACAGACTC 1575
DB 495 AsnProAsnGlyGlnGlyLeuProHisTrpProGlnTyrlleuArgGlnGlnGlnGlyLeu 514
QY 1576 CAGCTGACCTTGACATGAGCTCGGACAGAGACTCAAGAAACCGCGGAGTGAATTTTGG 1635
DB 515 GlnIleGlyAlaThrThrGlnGlnAlaGlnArgLeuLysAlaGlnGlyValAlaIlePheTrp 534

```

```

QY 1636 ACCAGCACCATC 1647
DB 535 ThrGlnLeuLeu 538

RESULT 7
S19307
N:Altermate names: proline-beta-naphthylamide
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
R:Matushima, M.; Inoue, H.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Takahashi
PDBS Letc. 293, 37-41, 1991
A:Title: The nucleotide and deduced amino acid sequences of porcine liver proline-beta-n
A:Reference number: S19307; MUID:92070571; PMID:1959668
A:Accession: S19307
A:Molecule type: mRNA
A:Residues: 1-566 <MAT1>
A:Cross-references: UNIPROT:Q29550; EMBL:X63323; NID:g1930; PIDN:CAA44929.1; PID:g1931
A:Accession: S23607
A:Molecule type: protein
A:Residues: 19-40 <MAT2>
A>Note: 28-Lys and 33-Leu were also found
C:Superfamily: Cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
P:1-18/Domains: signal sequence #status predicted <SIG>
P:19-566/Product: carboxylesterase #status experimental <MAT>
P:51-552/Domains: cholinesterase homology <CH>
P:80/Binding site: carbonyl (Asn) (covalent) #status predicted
P:222,467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 6,98e-82 Length: 566
Score: 1162.50 Matches: 250
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 44.33% Mismatches: 202
Query Match: 36.67% Indels: 41
DB: 2 Gaps: 8

US-10-023-515-3 (1-1746) x S19307 (1-566)
QY 34 TGTGCTTTTCTGATTTCTCCAGCCCTGTGGGACACAGACAGTGGGAAAACTGGG 93
DB 2 TrpLeuLeuProLeuValLeuThrSerLeuAlaThrTrp-----AlaGly 19
QY 94 CTTTGTCTGAAGGCGCACAGAGAAACACAGGCTGGATTCAGGCGCAAGCATC 153
DB 20 GlnProAlaSerProProValAlaAspThrAlaGlnGlyArgValLeuGlyLysTyrlle 39
QY 154 ACTGTCGCGGAGAGCCCTGTGCTGTGAACGTTCTCGAGTCCCTTGTGCTGCTGCC 213
DB 40 SerLeuGlnGlyLeuAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 59
QY 214 CCGCTGGATCCCTCGATTAACAAACCGGACCGCTGATCGCCCTGGGATTAATCGCA 273
DB 60 ProLeuGlnSerLeuArgPheAlaProProGlnProAlaGlnProTrpSerPheValLys 79
QY 274 GAAGCACCCTCTACCTTAATTGTGCTCCAGAAC-----TCA 312
DB 80 AsnThrThrSerTyrlleProMetCysGlnAspProValAlaGlnGlnMetThrSer 99
QY 313 GAGTGGCTGCTTATGATCAACACATGCTCAAGTGCATTACCCGAAATTCGAGTGCTA 372
DB 100 AspLeuPheThrAsnGlyLysGlnArgLeuThrLeuGlnPhe-----Ser 114
QY 373 GAAGACTGCTTACTTAACAATCTATGCGCTGCGCAACCGCATACAGGCTCAAGCTC 432
DB 115 GlnAspCysLeuTyrlleuAsnIleTyrlleProAlaAspLeuThrLysArgGlyArgLeu 134
QY 433 CCGTCTTGTGTGTTTCCAGAGAGTGCCTTCAAGACTGCTCAGCTCCATCTTGTAT 492
DB 135 ProValMetValTrpIleHisGlyGlyLeuValLeuGlyGlyAlaProMetTyrlleAsp 154

```

```

QY 493 GGGTCGCGCCCTGGCTGCTATGAGAGACCTGCTGGTGGTCTGCTCAAGTACCGGCTAGAGA 552
DB 155 G|yValValLeuValAlaHnIseG|uAenValValValAlaAlaIleG|nTrpHrG|uLeuG|y
QY 553 AATATTGGTTTCTTCAACCATGGGATCAAGCATGCTCCGGGAACTGGGCTTTCAAGAC 612
DB 175 I|eTrpG|yPhaPheSerThrG|yAerG|uHnIseSerA|yAenTrpG|yHnIseAerP 194
QY 613 CAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
DB 195 G|nValAlaAlaLeuHnIseTrpValG|nG|uAenHnIleAlaAenPheG|yG|yAerProG|y 214
QY 673 TCTGTGACCATGTTTGGCGAGTCCGGCGAGCATTAAGTGTTCATGCTTAACTGCTCT 732
DB 215 SerValThrIlePheG|yG|uSerAlaG|yG|uSerValSerValLeuValLeuSer 234
QY 733 CCCATGGCCCAAGGCTTATTCACAAAGCATCATGAGAGATGGGGTGGCCATCATCCCT 792
DB 235 P|roLeuAlaLysAenLeuPheHnIseA|yAlaIleSerG|uSerG|yValAlaLeuThrVal 254
QY 793 TACCTGAGAGCCCATGATTAAGAGAAAGTGAAGACCTGAGTGGTGGCAATTTCTGT 852
DB 255 A|aLeuValAlaG|yAerPheMetLysAlaAla-----A|aLysG|nIleAlaValLeuAla 272
QY 853 GGTAACAAATGCGTCAAGACTCTGAGGCGCTGAGGCTGCTGAGAACAAACCTTCAAG 912
DB 273 G|yCyLysThrThrThrSerAlaValPheValHnIseCyLeuA|yG|nLysSerG|uAer 292
QY 913 GAGCTGTCGACCCCTGAGCCAGAAACAAAGCTTTCACCT----- 951
DB 293 G|uLeuLeuAerPheThrLeuLysSerLysPheLeuThrLeuAerPheHnIseG|yAerG|n 312
QY 952 -----CGAGTGGTTGATGGTCTTCTTCTTCAATGAGCCT 987
DB 313 A|yG|uSerHnIaerPheLeuProThrValAlaAerG|yValLeuLeuProLysMetPro 332
QY 988 CTAGATCTATTGTTCTCAAGAAAGCATTTAAAGCATTTCTTCATCATGAGTCAATTAAC 1047
DB 333 G|uG|uIleLeuAlaG|uLysAerPheAenThrValProG|yAlaValG|yIleAenLys 352
QY 1048 CACGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
DB 353 G|nG|uPheG|yTrpLeuLeuProThrMetMetG|yPheProLeuSerG|uLysLeu 372
QY 1105 AACAAAGTCCCTGGCCCTCATCTGATCAAAAC-----ATCCTGACATCCCGCCT 1155
DB 373 A|yG|nLysThrAlaThrSerLeuLeuTrpLysSerYrProIleAlaAenIleProG|u 392
QY 1156 CAGTATTGGACCTTGGCTTAATGAATACTTCATGACAAAGACCTCCCTGACCTGAATTC 1215
DB 393 G|uLeuThrProValAlaThrAerPheLysTrpLeuG|yG|yTrpAerAerProValLysLys 412
QY 1216 CGAGACAGCTTTTGGAGTCTGCTGAGAGATGTTCTTGGTCCCGCTGACATGACA 1275
DB 413 LysAerPheLeuPheLeuAerPheMetG|yAerValValPheG|yValProSerValThrVal 432
QY 1276 GCTCGATATCAACAGATGCTGTCGACCTGTCTAATTCTTAAGATTCCGACCGGCT 1335
DB 433 A|aA|yG|nHnIseAerPheG|yAlaAerThrTrpMetYrG|uPheG|nTrpA|yAerPro 452
QY 1336 CAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGC 1395
DB 453 SerPheSerSerAerLysLysProLysThrValIleG|yAerPheG|yLysAerG|nIlePhe 472
QY 1396 TTTGTGTGGTGGTGGCTTCTGTAAGGGGAGCATTTGTTATGTTGAAGAGCCACGGAG 1455
DB 473 SerValPheG|yPheProLeuLeuLysAerPhe-----A|aProG|u 486
QY 1456 GAGGAGAGATTACTAGCCGGAAGATGTAATAACTGGGCTTACCTTGTCTGAACCGGG 1515
DB 487 G|uG|uIValaSerLeuSerLysThrValMetLysPheTrpAlaAenPheAlaAerSerG|y 506
QY 1516 AATCTTAATGGAAAGACCTGTCTGTGGCCAGCTTAATTAATCTGACTGACAGTACCTC 1575

```

```

DB 507 AenProAenG|yG|uLysG|yLeuProHnIseTrpProMetYrAerPheG|nG|uLysG|yTrpLeu 526
QY 1576 CAGCTGACCTTGAACATGAGCCTCGACAGAGACTCAAGAAACCGGGTGGATTTTGG 1635
DB 527 G|nIleG|yValAenThrG|nAlaAlaLysAerG|uLysG|yG|uValAlaAerThrP 546
QY 1636 ACCAGCACCATC 1647
DB 547 AenAerPheLeu 550

RESULT 8
JX0054
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Aug-1998
C:Accession: JX0054
R:Takegi, Y.; Morohashi, K.; Kawabata, S.; Go, M.; Omura, T.
J. Biochem. 104, 801-806, 1988
A:Title: Molecular cloning and nucleotide sequence of cDNA of microsomal carboxylesterase
A:Reference number: JX0054; MUID:89174514; PMID:3235453
A:Accession: JX0054
A:Molecule type: mRNA
A:Residues: 1-549 <TAK>
A:Experimental source: Liver
C:Superfamily: cholinesterase; cholinesterase; glycoprotein; mitochondrion
C:Keywords: carboxylic ester hydrolase; cholinesterase; glycoprotein; mitochondrion
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-548/Product: carboxylesterase E1 #status predicted <MAT>
F:50-538/Domain: cholinesterase homology <CHN>
F:79,274,302,375,476/Binding site: carbonyl site: Ser, His #status predicted
F:221,453/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,88e-80 Length: 549
Score: 1144.00 Matches: 238
Percent Similarity: 60.33% Conservative: 92
Best Local Similarity: 43.51% Mismatches: 183
Query Match: 36.09% Indels: 34
DB: 2 Gaps: 9

US-10-023-515-3 (1-1746) x JX0054 (1-549)
QY 46 CTGATCTCCAGCCCTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGTA 105
DB 11 LeuAlaValCySProIleTrpG|yHnIse-----ProSerSerPro 23
QY 106 GGGCCACAGAGAAACACACAGCTGGGATTCAGGCAAGCAAGTCACTGTCTGGGA 165
DB 24 ---ProValAlaPheThrThrLysG|yLysValLeuG|yLysTrpValSerLeuG|uG|y 42
QY 166 AGCCCTGTGCTGGAACGTTCTCTGGAGTCCCTTGTCTGCTCCCGCTGGATTC 225
DB 43 PheThrG|nProValAlaValPheLeuG|yValProPheAlaLysProProLeuG|ySer 62
QY 226 CTGGCATTTTCAACACCGGACCTGTCATCGCCCTGGGATTAATCTGCAAGAACCACTCC 285
DB 63 LeuA|yG|nHnIaerProG|uProAlaG|uP|roTrpSerPheValLysAenThrThr 82
QY 286 TACCTTAATTGGCTTCCGAAC-----TCAGTGGCTGCTC 324
DB 83 TyrProMetCySPheG|nAerG|yValValG|yLysLeuLeuAlaAerMetLeuSer 102
QY 325 TTGATCAACACATGCTCAAGTGCATTAACCGGAATTCGATGTCAGAAAGTCTGCTC 384
DB 103 ThrG|yLysG|uSerIleProLeuG|uPhe-----SerG|uAerCyPhe 117
QY 385 TACCTGAACATCTATGCGCTGCTGCCAGCGCATCAAGCTCCAGCTCCGCTTGTG 444
DB 118 TyrLeuAenIleTrpSerProAlaAerPheLeuThrLysAenSerTrpLeuProValMetVal 137
QY 445 TGCTTCCAGAGAGTGGCTTCAAGACTGCTCAAGCTTCACTTTGATGGGCTCGCCCTG 504

```

```

Db      138  TrrlrlslgllyglyleullellellyglyalaserProtyrserglyleuallaleu 157
Oy      505  GCTGCTATGAGAGCGTGGTGGTGGTCCGTCAGTACCGGCTGAGAAATATGGTTTC 564
Db      158  Serlslslglsluaslvalvalvalthrlllelgllyrhlrgllylterpdyneu 177
Oy      565  TTCAACCATGCGGATCAGATGCTCCGGGAATCGGGCTTCAGAGACAGAGTGGCTCT 624
Db      178  PheSerthrglyasgluhslerarglyasenthrpallslleuaslgluaslala 197
Oy      625  CTGTCCTGGGTCCAGAGAAACATCGAGTTCTTGGTGGGACCCCAAGCTCTGTGACATC 684
Db      198  Leuargtrvalglaslslaslallelslaslsllelslaslslaslslaslslsl 217
Oy      685  TTGGCGAGTCCGGGGAGCCATAGTGTCTTACTCTTATCTGCTGCTCCATGGCCAA 744
Db      218  PhegllysluSerlslgllyslvalserlslalalevalleuSerProeuallay 237
Oy      745  GCGTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCCCTTACCTGAGGCC 804
Db      238  AsneuPhehlslarglslallelsergluSerglyvalleuThrThrAsneuasl 257
Oy      805  CATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 864
Db      258  Lysaslslslslslslslslslslslslslslslslslslslslslslslslsl 275
Oy      865  TCAGACTGAGAGCGCTGAGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
Db      276  Thrserlslslslslslslslslslslslslslslslslslslslslslslsl 295
Oy      925  CTCAGCCGAGAA-----ACAAAGCTTTCATCGAGTGGTGGTGGTGGTGGTGGT 978
Db      296  Leuthrvallyslslaslslslslslslslslslslslslslslslslslslsl 315
Oy      979  AATGAGCTTATGATCTATGCTCTGAGAGAGATTTAAAGCAATTCCTTCATCGAG 1038
Db      316  LysThrProgluGlulLeuThrGlulysSerPheaslslslslslslslslslsl 335
Oy      1039  GTCAATACACAGAGTGGCTCTGCTGCTGCTTATGAGAGAGAGAGAGAGAGAGAG 1098
Db      336  Pheaslslslslslslslslslslslslslslslslslslslslslslslslsl 355
Oy      1099  GGC---TCACACAGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
Db      356  GlyargmetaslglulylmetaslaserPheleuylslslslslslslslslslsl 375
Oy      1147  ATCCCGCTCATGATTTGACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206
Db      376  llesergluSerValllleProalllellelglulylslslslslslslslslsl 395
Oy      1207  ACTGAAATCCAGACAGTCTTGTGAGCTTGTGAGAGATGTTTGTGTGCTCCCTGCA 1266
Db      396  Alslslslslslslslslslslslslslslslslslslslslslslslslslsl 415
Oy      1267  CTGATCAGAGCTGATCATCAGAGATGCTGGTGGACCTGCTGCTGCTGCTGCTGCT 1326
Db      416  ValleuSerSerhlslslslslslslslslslslslslslslslslslslslsl 435
Oy      1327  CACCGGCTCATGCTTGTGAAGACAGAGCGGCTTTTGTCAAGCGGACCGACCTGAT 1386
Db      436  Tyrhlslslslslslslslslslslslslslslslslslslslslslslslsl 455
Oy      1387  GAATGCGCTTGTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1446
Db      456  Glulslslslslslslslslslslslslslslslslslslslslslslslslsl 469
Oy      1447  GCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
Db      470  AlslsergluGlulThrAsneuSerlyslslslslslslslslslslslslslsl 489
Oy      1507  CGAACCGGAGATCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1566
Db      490  Argaslslslslslslslslslslslslslslslslslslslslslslslslsl 509

```

```

Oy      1567  CAGTACTCCAGCTGAGTGAATGACATGAGCCTCGAGACAGACTCAAGAACCGGGGTG 1626
Db      510  GlyTyrleuGlulsllellyslalThrThrGlulslslslslslslslslslslsl 529
Oy      1627  GATTTTGGACGACGACCATC 1647
Db      530  Alslslslslslslslslslslslslslslslslslslslslslslslslslsl 536

RESULT 9
A1010
A:Alveolar lesterase (BC 3.1.1.1) precursor, monocyte/macrophage [validated] - human
N:Alternate names: carboxylesterase, hepatic; monocyte/macrophage serine esterase
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
R:Accession: A1010; JH0327; K47376; A49816; PS0280; 161085; A48809; I57004
J:Biol. Chem. 266, 18832-18838, 1991
A:Title: A serine esterase released by human alveolar macrophages is closely related to )
A:Reference number: A41010; MUID:92011649; PMID:1918003
A:Accession: A41010
A:Molecule type: mRNA
A:Residues: 1-567 <MUN>
A:Cross-references: UNIPROT:P23141; GB:M73499; NID:G179927; PIDN:AAA35649.1; PID:G179928
A:Note: parts of this sequence, including the amino end of the mature protein, were confi
R:Long, R.M.; Calabrese, M.R.; Martin, B.M.; Pohl, L.R.
Life Sci. 48, PL43-PL49, 1991
A:Title: Cloning and sequencing of a human liver carboxylesterase isoenzyme.
A:Reference number: JH0327; MUID:91148424; PMID:1997784
A:Accession: JH0327
A:Molecule type: mRNA
A:Residues: 61-567 <LON>
A:Cross-references: GB:M55509; NID:G179929; PIDN:AAA35650.1; PID:G179930
R:Shibata, F.; Takagi, Y.; Kitajima, M.; Kuroda, T.; Omura, T.
Genomics 17, 76-82, 1993
A:Title: Molecular cloning and characterization of a human carboxylesterase gene.
A:Reference number: A47376; MUID:94010913; PMID:8406473
A:Accession: A47376
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA; DNA
A:Residues: 1-3, 'PALV', 8-11, 'A', 13-567 <SHI>
A:Cross-references: GB:D21089; NID:G455476; PIDN:BA04650.1; PID:G458470
A:Note: sequence extracted from NCBI backbone (NCBI:137630) and corrected to correspond
R:Zschunke, F.; Salmassi, A.; Kneipe, H.; Buck, F.; Parwaresch, M.R.; Radzun, H.J.
Blood 78, 506-512, 1991
A:Title: cDNA cloning and characterization of human monocyte/macrophage serine esterase-1
A:Reference number: A49816; MUID:91300111; PMID:2070086
A:Accession: A49816
A:Molecule type: mRNA
A:Residues: 'G', 65-185, 'G', 187-361, 363-567 <ZSC>
A:Cross-references: GB:X52973; NID:G36421; PIDN:CAA37147.1; PID:G1335304
R:Riddle, P.W.; Richard, L.J.; Bowles, M.R.; Pond, S.M.
Gene 108, 289-292, 1991
A:Title: Cloning and analysis of a cDNA encoding a human liver carboxylesterase.
A:Reference number: PS0280; MUID:92084150; PMID:1748313
A:Accession: PS0280
A:Molecule type: mRNA
A:Residues: 114, 'H', 116-280, 'A', 282-300, 'IGNSTLYTTRTQREST', 318-336, 'R', 338-382, 'GSPT', 31
A:Cross-references: GB:M65261; NID:G187028; PIDN:AAA83932.1; PID:G187029
A:Experimental source: liver
A:Note: differences between this sequence and other reports appear to be due to frameshift
R:Kroetz, D.L.; McBride, O.W.; Gonzalez, F.J.
Biochemistry 32, 11606-11617, 1993
A:Title: Glycosylation-dependent activity of baculovirus-expressed human liver carboxyle
A:Reference number: A48809; MUID:94032283; PMID:8218228
A:Accession: 161085
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-55, 'G', 57-361, 363-535, 'G', 537-567 <KRO1>
A:Cross-references: GB:L07765; NID:G180949; PIDN:AAA35711.1; PID:G180950
A:Accession: A48809
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```


A:Residues: 1-17, 'A', 18-55, 'G', 57-535, 'G', 537-567 <R02>
A:Cross-references: GB:L077664; NID:q180947; PIDN:AA16036.1; PID:q180948
A:Genetics:
A:Gene: GDB:CEB1; HMSE
A:Cross-references: GDB:128044; OMIM:114835
A:Map position: 16q13-16q22.1
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; liver
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-145/Domain: properdin #status predicted <PRO>
F:50-553/Domain: cholinesterase homology <CHE>
F:146-567/Product: carboxylesterase #status experimental <MAT>
F:564-567/Region: endoplasmic reticulum retention signal #status atypical
F:221,468/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	3,88e-80	Length:	567
Score:	1140.00	Matches:	248
Percent Similarity:	58.39%	Conservative:	79
Best Local Similarity:	44.29%	Mismatches:	197
Query Match:	35.96%	Indels:	36
DB:	1	Gaps:	9

US-10-023-515-3 (1-1746) x AA1010 (1-567)

```
QY 34 TGGTCTTTTTCCTGATTCTCCAGCCCTGTGGGACACAGAGGGGAAAACCTGGG 93
DB 2 TTPLEUAGLALPHALIELEUAlALHREUSERAlASERAlALATRPGLYHIS----- 19
QY 94 CTTTCTGTGAAGGGCCACAGAGAACACAGCTGGAGTGTTCAGGGCAAGCAAGTC 153
DB 20 ProSerSerPro---ProValValAspThrValHISGLYValLeuGLYLeuPheVal 38
QY 154 ACTGTCGGGAGAGCCCTGTCTGTGAACGTTCCTCGAGTCCCTTGTGCTCTCC 213
DB 39 SerLeuGLYLeuPheAlGlnProValAlAlPheLeuGLYLeuProPheAlAspPro 58
QY 214 CCGTGGATCCCTGCGATTTACGAACCCGAGCTGCGATCGCCCTGGATTAACCTGCA 273
DB 59 ProLeuGLYProLeuArgPheThrProGlnProAlAGluProTrpSerPheValys 78
QY 274 GAAGCCACTCTCAACCTTAATTGCTGCCAAGACTCAAG-----TGCTGCTTTA 327
DB 79 AsnAlaThrSerTYrProPrometCYsrHrglnAspProLYsAlAGLYGlnLeuSer 98
QY 328 GATCAACACATGTCGTCAAGTGATTAACCGAAATTCCGAGTGTCAAGACCTGCTAC 387
DB 99 GluLeuPheThrAsnArgLYsGLUAsnAlleProLeuLYsLeuSerGLUAspCYsLeuTYr 118
QY 388 CTGAACATCTATGGCGCTGCCACGCGCATACAGGCTTCAAGCTCCCGCTTGGTGG 447
DB 119 LeuAsnAlleTYrThrProAlAspLeuThrLYsLYsAsnArgLeuProValMetValTrp 138
QY 448 TTCCAGAGAGTGCCTTCAAGATGCGCTCAAGCTTCCATCTTATGGTCCGCCCTGCT 507
DB 139 IleHISGLYLeuLeuMetValGLYAlAlAspThrTYrAspGLYLeuAlAlLeuAla 158
QY 508 GCGTATGAGACGTGCTGTGCTGTCAGTACCGGCTAGAGAAATTTGTTGTTCTTC 567
DB 159 AlAlHISGLUAsnValValValThrIleGlnTYrArgLeuGLYLeuTrpGLYPhePhe 178
QY 568 ACCATGAGGATCAGCATGCTCCGGGGAACCTGGACCTTCAAGACCAAGTGGCTGCTG 627
DB 179 SerThrGLYAspArgLHISerArgLYsAsnTrpGLYHISLeuAspArgLHISAlAlLeu 198
QY 628 TCCGGGTCCAGAGACATGAGTTCCTGGTGGGAGACCCACCTCTGTACCATCTTT 687
DB 199 ArgTrpAlGlnAspAsnAlleAlAspPheGLYLeuAsnProGLYSerValThrIlePhe 218
QY 688 GGCAGTCCGGGAGGACCATAGTTCCTTACTGCTTACTGCTCCCATGGCAAGGC 747
DB 219 GLYLeuSerAlAGLYLeuGLYLeuSerValSerValLeuValLeuSerProLeuAlLYsAsn 238
```

```
QY 748 TTATTCACAAAGCATCATGAGATGGGTGGCCATCATCCCTTACTGTAGAGCCAT 807
DB 239 LeuPheHISArgAlAlleSerGLYLeuSerGLYAlAlAlLeuThrSerValLeuValLYs 258
QY 808 GATTATGAGAAAG-----AGTGAAGACCTGACGGGTGTGACATTTCTGTATCAAT 861
DB 259 GLYAspValLYsProLeuAlAGLAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 275
QY 862 GCGTCAGACTGTGAGGCCCTGTGAGTGTGCTGAGGACAAACCTCCAGAGACTGTG 921
DB 276 ThrThrThrSerAlAlAlAlMetValHISCYsLeuArgLHISThrGLYLeuGLYLeu 295
QY 922 ACCCTGAGCCAGAAACAAAGCTTTTCACT----- 951
DB 296 GluThrThrLeuLYsMetLYsPheLeuSerLeuAspLeuGlnLYsAspProArgLys 315
QY 952 -----CGAGTGTGATGATGCTTTCTTCTTAATGAGCCCTGATCTA 996
DB 316 GlnProLeuLeuGLYThrValIleAspGLYMetLeuLeuLeuLYsThrProGLYLeu 335
QY 997 TTGTCTCAAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAGAGTGT 1056
DB 336 GlnAlAGLArgAsnPheHISThrValProLYsMetValGLYIleAsnLYsGlnGLY 355
QY 1057 GCGTTCCTGCTGCTTATGAGAGAGGCTCTGAGATCTCGAGTGGCTCC-----AACAG 1110
DB 356 GLYTrpLeuIleProMetGlnLeuMetSerTYrProLeuSerGLYLeuLeuAspArg 375
QY 1111 TCCCTTCCCTCCATCTGATCAAAAC-----ATCCGACATCCCGCTCACTAT 1161
DB 376 LysThrAlAlMetSerLeuLeuTrpLYsSerTYrProLeuValCYsIleAlAlLYsGlu 395
QY 1162 TTGACCTTGGGTATGAAATCTTCATGACACAGACTCCCTGACTGAAATCCAGAC 1221
DB 396 IleProGLYAlAlThrGLYLYsTYrLeuGLYLYsThrAspThrValLYsLYsAsp 415
QY 1222 AGTCTTGTGACTGTGATGAGATGTGTCTTGTGTGCTGCTGACATGATCACAGTCA 1281
DB 416 LeuPheLeuAspLeuIleAlAlAspValMetPheGLYAlProSerValIleValAlArg 435
QY 1282 TATACAGAGATGTGTGCACTGTCTACTTCTATGATTTTGCGACCGGCTCAAGTC 1341
DB 436 AsnHISArgAspAlAGLYAlAlProThrTYrMetTYrGlnPheGlnTYrArgProSerPhe 455
QY 1342 TTGAAGACAGAGAGCGGCTTTTGTCAAGCGACGACGCTATGAGTACGCCCTTGTG 1401
DB 456 SerSerAspMetLYsProLYsThrValIleGLYAspHISGLYAspGLYLeuPheSerVal 475
QY 1402 TTGGTGTGCTCTTCTGAGAGGGGACATTTATGTTTCGAAGAGCCACGAGAGAG 1461
DB 476 PheGLYAlAlProPheLeuLYs-----GluGLYAlAspArgLHISGLYLeu 489
QY 1462 AAGTTACTGACCGGAAAGATGATGAATAACTGGGCTACCTTGTCTGAAACGGGAATCT 1521
DB 490 IleArgLeuSerLYsMetValMetLYsPheTrpAlAlAsnPheAlAlArgAsnGLYAsnPro 509
QY 1522 AATGGAAACAGACCTGTCTGTGCGCAGCTTATTAATCTGACTGAGAGTACCTCCAGCT 1581
DB 510 AsnGLYLeuGLYLeuProHISThrProGLYTYrAsnGlnLYsGLYLeuGLYLeuGlnIle 529
QY 1582 GACTTGAACATGAGCCTCGGACAGAGACTCAAGAACCGGCGGTGATTTTGGACAGC 1641
DB 530 GLYAlAlAsnThrGlnAlAlAGLInLYsLeuLYsAspLYsGLYAlAlAlPheTrpThrAsn 549
```

RESULT 10

62788 carboxylesterase (EC 3.1.1.1) BS-4 precursor, liver - rat
N:Alternate names: hydrolase B
C:Species: Rattus norvegicus (Norway rat)
C:date: 28-Oct-1996 #sequence revision 13-Mar-1997 #ext_change 09-Jul-2004
C:Accession: S62788; S51203; A55304; S49257
R:Robb, M.; van Schaftingen, E.; Beaufay, H.
Biochem. J. 313, 821-826, 1996

A:Title: Cloning and sequencing of rat liver carboxylesterase ES-4 (microsomal palmitoyl
A:Reference number: S62788; MUID:96190723; PMID:8611161
A:Accession: S62788
A:Molecule type: mRNA
A:Residues: 1-561 <ROB>
A:Cross-references: UNIPROT:064573; EMBL:X81825; NID:G550417; PIDN:CAA57419.1; PID:G5504
A:Experimental source: liver
R:Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A:Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID:7986098
A:Accession: S51203
A:Molecule type: protein
A:Residues: 19-48 <MOR>
A:Experimental source: liver
R:Yan, B.; Yang, D.; Brady, M.; Parkinson, A.
J. Biol. Chem. 269, 23688-23696, 1994
A:Title: Rat kidney carboxylesterase. Cloning, sequencing, cellular localization, and re
A:Reference number: A55304; MUID:95050819; PMID:7961958
A:Accession: A55304
A:Molecule type: mRNA
A:Residues: 1-6, 'P', 8-59, 'P', 61-212, 'A', 214-252, 'T', 254-309, 'IT', 312-341, 'N', 343-424, 'P'
A:Cross-references: GS:U010697; NID:G562007
A:Note: the sequence in GenBank entry RNU010697, release 107, (PID:G562008) has the codon
R:Robbi, M.; Beaufay, H.
submitted to the EMBL Data Library, September 1994
A:Reference number: S49257
A:Accession: S49257
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <RO2>
A:Cross-references: EMBL:X81825; NID:G550417; PIDN:CAA57419.1; PID:G550418
C:Function: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
A:Description: cholinesterase; cholinesterase homology
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-561/Product: carboxylesterase ES-4 #status experimental <MAT>
F:50-551/Domain: cholinesterase homology <CHE>
F:221,466/Active site: Ser, His #status predicted
F:301/Binding site: carboxylate (Aen) (covalent) #status predicted

Alignment Scores:
Pred. No.: 4,62e-80 Length: 561
Score: 1139.00 Matches: 243
Percent Similarity: 58.04% Conservative: 82
Best Local Similarity: 43.39% Mismatches: 195
Query Match: 35.93% Indels: 40
DB: 2 Gaps: 10

US-10-023-515-3 (1-1746) x S62788 (1-561)

```

QY 43 TTCTGATTTCTCCAGCCCTGTTGGAGACACAGAGTGGGAAAACTGGGCTTCTGCT 102
DB 5 PheLeuIleLeuValSerLeuAlaThrCysValValTyrGly-----AsnProSerSer 22
QY 103 GAGGCGCACAGAGAACACAGCGCTGGATGATTCAGGCGACAGCAATCACTGCTCTG 162
DB 23 Pro--ProValValAspThrThrTyrSgIlySValLeuGlyLyrValSerLeu 41
QY 163 GGAAGCCCTGGCGTGAAGGTTCCTCGGAGTCCCTTGTGCTGCTCCCGCGTGGGA 222
DB 42 GlyValThrGlnSerValAlaValPheLeuGlyValProPheAlaLysProLeuGly 61
QY 223 TCCCTGCATTTACGAAACCCGAGCTGATCGCCCTGGATATCTGGAGAACCAAC 282
DB 62 SerLeuAlaGlnPheAlaProProGlnProAlaGlnProTyrSerPheValLysAsnThr 81
QY 283 TCTTACCTTAATTTGTGCTCCAGAACTCAAGTGG-----CTGCTCTTA 327
DB 82 ThrTyrProProMetCysSerGlnAspAlaValSgIlyGlnArgMetAsnAspLeu 101
QY 328 GATCAACAGATGCTCAAGGTGATTAACCGAAATTCGAGGTGTCAAGAGATGCGCTTAC 387

```

```

DB 102 ThrAsnArgLysGlyLysIleHis-----LeuGlnPheSerGlyAspCysLeuTyr 118
QY 368 CTGAACATCTATGCGCCCTCCACCCGATACAGGCTCAAGGCTCCGCTTGGTGG 447
DB 119 LeuAlaIleTyrThrProAlaAspPheThrLysAsnSerAspGlyLeuProValTyr 138
QY 448 TTCCAGAGAGGCGCTTCAAGACTGCTCAGGCTCATCTTGTGATGGGTCCGCGTGC 507
DB 139 IleHisGlyGlyGlyMetThrLeuGlyValAsnThrTyrAspGlyArgValLeuSer 158
QY 508 GCTTATAGAGAGTGTGCTGTGCTCTGATGACGAGCTTGAATATTTGTTTCTTC 567
DB 159 AlaTyrGlyAsnValValValAlaIleGlnTyrArgLeuGlyIleTyrPhePhe 178
QY 568 ACCAATAGGATCAGCATGCTCCGCGGAACTGGGCTTCAAGAGACAGTGTGCTCTG 627
DB 179 SerThrGlyAspGlnHisSerArgLysAsnTyrGlyHisValAspGlnValAlaLeu 198
QY 628 TCCGTGGTCCAGAAAGATCGAGTTCTTGGTGGGACCCGAGCTCTGTGACATCTT 687
DB 199 HisThrValGlnAspAsnIleAlaAsnPheGlyAspProGlySerValThrIlePhe 218
QY 688 GCGAGTCCGCGGAGCCATAAGTGTCTTACTTATAGTGTCTCCATGGCCAAAGGC 747
DB 219 GlyLysSerAlaGlyGlyPheSerValSerValLeuValLeuSerProLeuThrLysAsn 238
QY 748 TTATTTCCAAAGCATCATGAGAGTGGGAGCCATCATCTTCACTGAGGCCCAT 807
DB 239 LeuPheHisArgAlaIleSerGlnSerGlyValValPheLeuProGlyLeuLeuThrLys 258
QY 808 GATTATGAGAGAGAGAGACCTCGAGGTGGTGGACATTTCTGTGTAACAATCGCTCA 867
DB 259 AspValAlaGlyProAla-----AlaLysGlnIleAlaAspMetAlaGlyCysGluThrThr 276
QY 868 GACTTGAAGCCCTGCTGAGTGTGCTGAGCAAAACCTCCMAAGAGCTGTGACCTTC 927
DB 277 ThrSerAlaIleIleValHisCysValLeuArgGlnTyrGlnGluGlnLeuGluIle 296
QY 928 AGCCGAAACAAAG----- 942
DB 297 MetLysLysMetAsnLeuIleLysLeuSerGlnArgAspAsnLysGlySerTyrHis 316
QY 943 TCTTCACTCGAGGTGATGAGTGTCTTCTTCTTAATGAGCTCTGATATATGCTT 1002
DB 317 PheLeuSerThrValValAspAsnValValLeuProLysAspProLysGlyIleValAla 336
QY 1003 CAGAAAGCATTTAAAGCAATCTTCCATCATCGGAGTCAATTAACAGAGTGGCTTC 1062
DB 337 GlyLysAsnPheAsnThrValProTyrIleValGlyIleAsnLysGlnGlySerTyr 356
QY 1063 CTGCTGCTT-----ATGAAGAGGCTCTGAGATCTTCAGTGGCTCCAAAGTCCCTT 1116
DB 357 LeuLeuProThrMetMetGlyPheValProAlaAspValGlu--LeuAspLysMet 375
QY 1117 GCCCTCCATCGATACAA-----AACATCCGACATCCGCGCTCAGATATTGGAC 1167
DB 376 AlaIleThrLeuLeuGlnLysPheAlaSerLeuTyrGlyIleProGlnAspIleIlePro 395
QY 1168 CTGTGGCTAATGATATCTTCATGACAGACACTCCCTGACTGACAAATCCGAGACGTCT 1227
DB 396 ValAlaIleGlyLysTyrArgLysSerAspAspSerIleLysIleArgAspGlyIle 415
QY 1228 CTGACCTGTGGAGATGTCTTGTGCTCTGATGATGATGATGATGATGATGATGAT 1287
DB 416 LeuAlaPheIleGlyAspValSerPheSerIleProSerValMetValSerArgAspHis 435
QY 1288 AGAGATGTGGTGGACCTGTACTTCTATGAGTTTGCGACCGGCTCAGAGCTTTGA 1347
DB 436 ArgAspAlaGlyAlaProThrTyrMetCysTyrGlnTyrGlnTyrTyrProSerPheSer 455
QY 1348 GACACGAAGCGGCTTTGTCAAGACCGACACGCTGATGAAGTCCGCTTGTGTGCT 1407

```

Dh 456 ProGlnAArgProLysHleValValGlyAerPHLeAAspAAspLeuTyrSerValPheGly 475
QY 1408 GGTGGCTTCTGTAAGGGGACATTTGTAAGTTCGAAAGGACCAAGGAGGAAAGTTA 1467
Dh 476 AlaProIleLeuArg-----AaPGlyAlaSerGlnGlnIleVal 489
QY 1468 CTGAGCCGGAAGATGATGAAATTAAGTCTGAGCTTGTGTCGAAACGGGAACTTAAGG 1527
Dh 490 LeuSerIlyMetValMetIlySerPheThrAlaAenPheAlaArgAAsnGlyAAsnProAAsnGly 509
QY 1528 AACGACCTGTCTGTGGCCAGCTTAATTAATGTAAGTGAAGCACTTCCAGCTGACCTTG 1587
Dh 510 ArgGlyLeuProHleThrProGlnTyrAerGlnTyrGlnGlnTyrLeuGlnIleGlyAla 529
QY 1588 AACATGACGCTCGACAGAGACTCAAGAAGACCGGGGTGGATTTTGGACCAAGCAATC 1647
Dh 530 ThrThrGlnIleSerGlnArgLeuLysAlaGlnIleValAlaPheThrThrGlnIleLeu 549
RESULT 11
A55281
carboxylesterase (BC 3.1.1.1) egasyn - mouse
N:Alternate names: beta-glucuronidase endoplasmic reticulum-targeting protein; esterase-
C:Species: Mus musculus (house mouse)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C:Accession: A55281
R:Ovnic, M.; Swank, R.T.; Fletcher, C.; Zhen, L.; Novak, E.K.; Baumann, H.; Heintz, N.;
Genomes 11, 956-967, 1991
A:Title: Characterization and functional expression of a cDNA encoding egasyn (esterase-
A:Reference number: A55281; MUID:92147141; PMID:1783403
A:Accession: A55281
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Reissues: 1-562 <OVN>
A:Cross-references: UNIPROT:064176; GB:S80191; NID:9244727; PIDN:AA821335.1; PID:9244728
A>Note: sequence extracted from NCBI backbone (NCBI:80191, NCBI:80194)
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum
F:51-552/Domain: cholinesterase homology <CHS>
F:222,467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 5,06e-80 Length: 562
Score: 1138.50 Matches: 240
Percent Similarity: 58.52% Conservative: 76
Best Local Similarity: 44.44% Mismatches: 195
Query Match: 35.91% Indels: 29
DB: 2 Gaps: 6

US-10-023-515-3 (1-1746) x A55281 (1-562)

QY 91 GGGCTTCTGCTGAAGGGCCACAGAGCAACAGGCTGGATGGATTCAAGGCAAGCA 150
Dh 19 GLyHisProSerSerProProMetValAAspThrValGlnGlyLysValLeuGlyLysTyr 38
QY 151 GTCACTGTGCTGGAGACGCTGTGCTGTAACGTGTTCTCGAGTCCCTTGTGCT 210
Dh 39 IleSerLeuGlnGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaLys 58
QY 211 CCCCCGTGGATCCCTCGGATTTTAAGAACCCGAGCTGCATGCGCTCGGGAATAC 270
Dh 59 ProProLeuGlnGlySerLeuArgPheAlaProProGlnProAlaGlnIleProTyrSerSerVal 78
QY 271 CGAAGAACCACTCTACCTTAATTTGTGCTCCAGAACTCAGAGTGAGCTCTTATGAT 330
Dh 79 LysAAsnAlaThrSerTyrProProMetCysPheGlnAAspProValThrGlnIleVal 98
QY 331 CAACACAGTCTCAAGGTGATTAACCCGAATTCGAAGT-----TCAGAAAGTGCCTC 384
Dh 99 AAsnAAspLeuLeuThrAAsnArgLysGlyLysIleProLeuGlnPheSerGlnAAspCysLeu 118
QY 385 TACCTGAACATCTATGGCCCTGCCACGAGCGCATACAGGCTCCAGCTCCCGTCTTGAT 444
Dh 119 TyrIleuAenIleTyrThrProAlaAspLeuThrLysSerAAspArgLeuProValMetVal 138

QY 445 TGGTCCCAAGAGGTCCTTCAAGACTGCTCAGCTTCCATCTTGAATGGGTCCGCTG 504
Dh 139 TrpIleHleGlyGlyGlyLeuValLeuGlyValAAserThrTyrAAspGlyLeuValLeu 158
QY 505 GCTGCTATGAGAGCGTGTGTTGGTGGCTCCAGTCCCGGCTAGGAATATTTGGTTTC 564
Dh 159 SerThrHleGlnAenValValValValValIleGlnTyrAAspLeuGlyIleTrpGlyPhe 178
QY 565 TTCACCAATGGATCAGACTGCTCCGGAACTGGGGCTTCAAGACCAAGTGGTGTCT 624
Dh 179 PheSerThrGlyAAspGlnHleSerArgLysAntTrpGlnHleAAspGlnValAlaAla 198
QY 625 CTGTCCGGGGCCCAAGAAATGATGATTTCTGCTGGGACCCCACTGTGTACCATC 684
Dh 199 LeuHleThrValGlnAAspAAsnIleAlaLysPheGlyGlyAAspProGlySerValThrIle 218
QY 685 TTTGGCAGTCCGGGGAGCCATAAGTGTCTTATGATGTGTCTCCATGGCCAA 744
Dh 219 PheGlyGlnSerAlaGlyGlyGlyGlnSerValSerValLeuValLeuSerProLeuAlaLys 238
QY 745 GGTATTTCACAAAGCCATGATGAGAGAGTGGGTGGCCATATCCCTTACCTGAGGCC 804
Dh 239 AAsnLeuPheGlnArgAlaIleSerGlnSerGlyValAlaLeuThrAlaGlyLeuValLys 258
QY 805 CATGATTATGAGAGAGTGAAGGACCTGCAGTGGTGCACATTTCTGTGTAACAATGCG 864
Dh 259 LysAAsnThrArgProLeuAlaGlnLysIleAlaValIleSerGlyCysLysAAsn----- 276
QY 865 TCAGACTCTGAGGCTCTGCTGAGTGGCTGAGGACAAACCCCTCCAGAGAGTGTGACC 924
Dh 277 ThrThrSerAlaAlaMetValHleCysLeuAArgGlnTyrThrGlnGlnIleLeuGly 296
QY 925 CTCAGCCCAAGAAACAAGTCTTTC----- 948
Dh 297 ThrThrLeuLysLeuAenLeuPheLysLeuAAspLeuHleGlyAAspSerArgGlnSerHle 316
QY 949 -----ACTGAGGTGTTGATGAGTGGCTTTCTTCTTATGAGCTTCAGATCTATTG 999
Dh 317 ProPheValProThrValLeuAAspGlyValLeuLeuProLysMetProGlnGlnIleLeu 336
QY 1000 TCTCAGAAAGCATTTTAAGCAATTCCTTCATCATCCGAGTCAATTAACCAAGTGTGCG 1059
Dh 337 AlaGlnLysAAsnPheAAsnThrValProTyrIleValGlyIleAAsnLysGlnIlePheGly 356
QY 1060 TTCTCTGCTGCT--ATGAAGAGGCTCTCGAGATCTCTCAGTGGCTCAACAAGTCCCTT 1116
Dh 357 TrpIleLeuProThrMetMetAAsnTyrProProSerAAspValLysLeuAAspGlnMetThr 376
QY 1117 GCCCTCATCTGATACAAAC-----ATCCTGACATCCCGGCTCAGTATTTCGAC 1167
Dh 377 AlaMetSerLeuLeuLysLysSerSerPheLeuLeuAAsnLeuProGlnAAspAlaIleAla 396
QY 1168 CTGTGGCTAATGATATCTTCATGACAAAGCACTCCCTGACCTGAATCCGAGACAGTCTT 1227
Dh 397 ValAlaIleGlnLysTyrIleAAspArgLysAspTyrThrGlyArgAAsnLysAAspGlnLeu 416
QY 1228 CTGACCTGCTTGAAGATGTGTTCTTTGTGCTCCCTGACATGATCACAGCTCGATAC 1287
Dh 417 LeuGlnLeuIleGlyAAspValAlaPheGlyValProSerValIleValSerAAspGlyHle 436
QY 1288 AGAGATGCTGGTGCACCTGTCTATCTTATGAGTTCGAGACCGGCTCAGTCTTGAA 1347
Dh 437 ArgAAspAlaGlyAlaProThrTyrMetTyrGlnPheGlnTyrSerProSerPheSerSer 456
QY 1348 GACACGAAGCCGGCTTTTGTCAAAGCCGACCAAGCTGATGAATCCGTTGTGTTCCGT 1407
Dh 457 GluMetLysProAAspThrValValGlyAAspHleGlyAAspIleLysTyrSerValPheGly 476
QY 1408 GGTGCTTCTGAAAGGGGACATTTGTTATGTTGTAAGAGCCACGAGAGGAAAGTTA 1467
Dh 477 AlaProIleLeuArgLys-----GlyThrSerGlnGlnIleAAsn 490


```

Db      481 ArgAsnGlyAsnProAsnGlyGluGlyLeuProHisTrpProIleTyrAspGlnGlySerGlu 500
QY      1567 CAGTACCTTCACGCTGAGCTTGAACATGAGCCTTCGACAGACAGCTCAAGAACCCGGGGTGT 1626
Db      501 GtATyTLeuGlnIleGlyAlaThrThrGlnGlnAlaGlnGlyLeuGlyGlyGlnIleVal 520
QY      1627 GATTTTGGACAGACAGCCATC 1647
Db      521 AlaPheTrpThrGlnLeuLeu 527

RESULT 13
S71597
carboxylesterase (EC 3.1.1.1) precursor, liver - rat
N.Alternate names: hydrolase C
C.Species: Rattus norvegicus (Norway rat)
C.Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 13-Sep-1998
C.Accession: S71597
Arch. Biochem. Biophys. 317, 222-234, 1995
A.Title: Cloning and expression of hydrolase C, a member of the rat carboxylesterase family
A.Reference number: S71597; MUID:95177656; PMID:7872788
A.Accession: S71597
A.Status: not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 1-561 <YAN>
A.Experimental source: liver; endoplasmic reticulum
C.Function:
A.Description: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
C.Keywords: carboxylic ester hydrolase; cholinesterase homology
F.1-18/Domain: signal sequence #status predicted <SIG>
F.19-551/Product: carboxylesterase #status predicted <MAT>
F.551-561/Domain: cholinesterase homology <CHS>
F.558-561/Region: endoplasmic reticulum retention signal
F.721,301/Binding site: carboxylate (Asn) (covalent) #status predicted
F.721,466/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 2,526-79 Length: 561
Score: 1,129.50 Matches: 236
Percent Similarity: 58.64% Conservative: 83
Best Local Similarity: 43.38% Mismatches: 188
Query Match: 35.63% Indels: 37
Gaps: 9
US-10-023-S15-3 (1-1746) x S71597 (1-561)

QY      91 GGGCTTGTGCTGAAGGGCCACAGACACAGAGCTGGATGATTCAGGGCAACAA 150
Db      18 GlyAsnProSerSerProProValAlaPheThrMetGlyGlyValLeuGlyLeuTyr 37
QY      151 GTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCCTCGAGTCCCTTGTGCTGT 210
Db      38 AlaSerLeuGlnGlyValThrGlnSerValAlaValPheLeuGlyValProPheAlaIys 57
QY      211 CCCCCGGGGGATCCCTGCGATTTTACGAACCGAGAGCTGATGATCCCTCGAGTAATCTG 270
Db      58 ProProLeuGlySerLeuArgPheAlaProProGlnProAlaGluProTyrSerPheVal 77
QY      271 CGAGAGCCACCTCTCAACCTTAATTTGTGTGCTCCAGAACTCAGAGTG----- 318
Db      78 LysAsnThrThrThrTyrProProMetCysSerGlnAspAlaThrIlyGlyGlnArgMet 97
QY      319 ---CTGCTCTTGAATCAACACATGCTCAAGTGCATTAACCGAAATTCGAGTGTCAAA 375
Db      98 AsnAspLeuLeuThrAsnArgGlyGlnIlyValHis-----LeuGlnPheSerGlu 114
QY      376 GACTGCTCTTACCTGAACATCTATGCGCTGCGCCACCGCCGATACAGGCTCCAAAGTCCC 435
Db      115 AspCysLeuTyrLeuAsnIleTyrThrProAlaPhePheThrIlyAspSerArgMetPro 134
QY      436 GTCTTGGTGTGTTCCAGAGAGTGCTTCAAGACGTGGCTCAGCTTCATCTTGTATGG 495

```

Db	135	ValMetValTprLlehiVgLyglYleuThrgInGlyAlaLaseThrTyraPrgLy	154
Qy	496	TCGCGCCCTGGCGCCCTATAGAGACGCTGCTGTGTGTCGTCACAGTACCGGCTAGAGATA	555
Db	155	AcgValIleuSerAlaTyrgIuamValValValAlaAlaIeGlnTyThrgIeuGlyIle	174
Qy	556	TTTGTGTTTTCACACATGGGATGACATGCTCCGGGGAACTGGGCTTCAAGACCG	615
Db	175	TpGlyPhePheSerThrgIyAspGlnHisSerArgGlyAsnTprGlyHisIleuAspGln	194
Qy	616	GTGGTGCCTCTCCCTGGGTCACAGAAACATTCGAGATTCTTGCGTGGGAGCCCACTCT	675
Db	195	ValAlaAlaIeunHisTprValGlnAspAsnIleAlaAsnPhcGlyGlyAspProGlySer	214
Qy	676	GTGACCAATTTTGGCGAGTCCGCGGAGCCATTAAGTGTTCATGTTTAACTGTCTCC	735
Db	215	ValThrIlePheGlyGlyUserAlaGlyGlyPheSerValSerValIleuValIleuSerPro	234
Qy	726	ATGGGCAAGGCTTATCCACAAGACGATCATGAGAGAGGGGATGGCCATCATCCCTTAC	795
Db	225	IeuSerIyAsnIeuTyHisArgAlaIleSerGlyIeuValValIleuIleThrgIn	254
Qy	796	CTGAGGCCCCATGATTAAGAGAAGAGGAGCATTCGAGCTGTGTCACATTTCTGTGT	855
Db	255	IeuPheThrIyAspValArgProAla-----AlaTyGlnIleAlaAspMetAlaGly	272
Qy	856	AACATGGGCTAGACTTCGAGGCCCTGCTGAGTGTGCTGAGACAAAACCTCCAGAG	915
Db	273	CysIyThrThrThSerAlaIleIleValHisCysIeuArgGlnIyThrgInGlnGln	292
Qy	916	CTGCTGACCCCTCAGCCACAGAA-----ACAAAG	942
Db	293	IeuIeuGlnIleMetGlnIyMetAsnIeuIeIyIeuSerSerGlnArgAspThrIyAs	312
Qy	943	TCT-----TTCACTCGAGTGTGATGTGCTTTCTTCTTAATGACCTCTTA	990
Db	313	GluSerTyHisAspIeuSerThrValIleAspAspValValIleuProIyAspProIyAs	332
Qy	991	GATCATTTGTCTCAGAAAGCATTTTAAACAAATCTTCATCATCATCGAGTCAATACAC	1055
Db	333	GluIleIeuAlaGlnIyAsnPheAsnThrValProTyIleValGlyIleAsnIyGln	352
Qy	1051	GAGTGTGGCTTCCTGCTGCT-----ATGAAGAGGCTCCTGAGATCCTCAAGTGGCTCC	1104
Db	353	GluCysGlyTprIleuIeuProThMetMetArgPheValProProAspValIyAs	371
Qy	1105	AACAGTCCCTGGCCCTCATCTGTATCA-----AACATGCTGACATCCCGGCT	1155
Db	372	AspIyIyMetAlaIleMetIeuIeuGlnIyAspIeAlaSerIleTyrgIyIleProGln	391
Qy	1156	CAGTATTTGACCTTGTGGCTAATGAATCTTCATGACAGACATCCCTGATCTGAATCT	1215
Db	392	AspIleIleProValAlaIleGlnIyTyArgIyAspIySerAspAspProIleIySerIle	411
Qy	1216	CGAAGACGCTTCCTGACCTGCTGAGATGTCCTTGTGTGTCCTCCGACATATAC	1275
Db	412	ArgAspGlyIleIeuAlaPheIleGlyAspValIleuPheCysIleIleProSerValMetVal	431
Qy	1276	GCTCGATATCACAGAGATGCTGTGGTGCACCTGTCTACTTCTATGATTTGGCACCGGCT	1335
Db	432	SerArgAspHisAspArgAlaGlyAlaIleProThrTyValTyGlnTyGlnTyIleTyArgPro	451
Qy	1336	CAGTGCCTTGAAGACAGAGCCGGCTTTTTCAAAGCCGACACGCTGATGAAGTCCGC	1395
Db	452	SerPheSerSerProGlnArgProIyAspValAlaGlyAspHisAlaAspAspValTyArg	471
Qy	1396	TTTGTGTTGGTGTGCTTCCTTAAGGGGACATGTGTTATGTCGAAGAGCCACGGAG	1455
Db	472	SerValPheGlyAlaProIleIeuArg-----AspGlyHisSerGln	485
Qy	1456	GAGGAGAAGTTACTGAGCCCGGAGAGATGAATACTGGGCTACCTTGTCTGAACCGGG	1515
Db	486	GlnGlnIleIeuIyIeuSerIyMetValMetCysIyThrIleAsnPheAlaArgAsnGly	505

```

QY 1516 AATCTAATGGAAGCAAGCTGCTCTGTCGACGCTTAATCTGACGTGACGACCTC 1575
DB 506 AsnProhAmlaAaGgLyLeuProHistrProGlnItyrAaPgluLygluIuLyrieu 525
QY 1576 CAGCTGACCTGAACATGAGCGCTCGACAGACACTCAAGAAACCGCGGTGATTTTGG 1635
DB 526 GlnIleGlyAlaThrThrlngInlserrGlnAgluLyAlaGluValAlaIaPheTrp 545
QY 1636 ACCGACCAATC 1647
DB 546 ThrGlnLeu 549

RESULT 14
JC2447
carboxylesterase (EC 3.1.1.1) ES-3 precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 29-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C/Accession: JC2447; S23462
R/Robbl. M.: Beaufay, H.
Biochem. Biophys. Res. Commun. 203, 1404-1411, 1994
A/Title: Cloning and sequencing of rat liver carboxylesterase ES-3 (EGASYN).
A/Reference number: JC2447; MUID:95032008; PMID:7945287
A/Accession: JC2447
A/Molecule type: mRNA
A/Residues: 1-561 <ROB>
A/Cross-references: UNIPROT:Q6J108; GB:X81395; NID:9550146; PIDN:CAA57158.1; PID:9550147
A/Experimental source: Liver
R/Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A/Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A/Reference number: S23460; MUID:9229908; PMID:1606962
A/Accession: S23462
A/Status: preliminary; translation not shown
A/Molecule type: mRNA
A/Residues: 503-554, 'R', 556-561 <MED>
A/Cross-references: EMBL:X65295; NID:957557; PIDN:CAA46390.1; PID:957558
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; glycoprotein
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-551/Product: carboxylesterase ES-3 #status predicted <MAT>
F/79,107,489/Binding site: carbohydrate (Aan) (covalent) #status predicted
F/221,466/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 8,04e-79 Length: 561
Score: 1123.00 Matches: 245
Percent Similarity: 57.07% Conservative: 86
Best Local Similarity: 42.24% Mismatches: 193
Query Match: 35.43% Indels: 56
DB: 2 Gaps: 11

US-10-023-515-3 (1-1746) x JC2447 (1-561)
QY 37 TGCCTTTCCGATTCCTCAGCCCTCTTGGACACAGACAGTGGGAAAAAATGGGCT 96
DB 2 CysLeuTyAlaLeuIleLeuValPheLeuAlaIaPheThrAlaGly-----GlyHis 19
QY 97 TCTCTGAAGGGCCACAGAGAACAGGCTGTGATGATTCAGGACAGCAAGTCACT 156
DB 20 ProSerSerLeuProValAlaPheThrLeuGlnGlyValValLeuGlyLysTrpValSer 39
QY 157 GTGCTGGAGACCCCTGTGCTGTGACAGCTGTCTCGAGAGTCCCTTTGGCTGCTCCCG 216
DB 40 LeuGlnGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaIaPhePro 59
QY 217 CTGGATCCCTCGGATTAACAACCGAGCTGTGATGGCCCTGGGATTAAGTGGAGAA 276
DB 60 LeuGlnSerLeuAaGpPheAlaProGlnProAlaGluProIlePheValIaAaPheVal 79
QY 277 GCACCTCTCAACCTTAATTTGTGCTCCAGAAC----- 309

```

```

DB 80 ThrThrSerTyProProMetCysSerGlnAspProValAlaGlyGlnIleValAsnAsp 99
QY 310 -----TCAGAGTGGCTGCTCTTAGATCAACAGACGTCAAGGTGCATTAACCGAAATTC 363
DB 100 LeuLeuThrAsnTrp-----GluGluAsnIleSerLeuGlnPhe----- 112
QY 364 GAGAGTGCAGAAAGCTGCTCTTACCTGAACATTAATGCGCTCCGACCGCATACAGGC 423
DB 113 -----SerGluAspCysLeuTyLeuAsnIleTyThrProAlaAspLeuThrLysArg 130
QY 424 TCCAAAGCTCCCGCTTTGGTGTGCTCCAGAGAGTGCCTTCAAGCTGCTCAACCTCC 483
DB 131 AspArgLeuProValMetValTrpIleHisGlyGlyGlyLeuValLeuGlyValAser 150
QY 484 ATCTTGATGGATCGCGCTGCTGCTCATGAGAGAGCTGCTGTGCTGCTCCAGTAC 543
DB 151 ThrTyAspGlyLeuAlaLeuSerThrHisGluAsnValValValValIleGlnTy 170
QY 544 CGGCTAGGAATATTTGGTTTCTTACCAACATGGATCAGATGCTCCGCGGAATCGGCGC 603
DB 171 ArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGluHisSerArgGlyAsnTrpGly 190
QY 604 TTCAAAGACCAAGTGGCTGCTGTCTGCTGCTCCAGAAAGACATGCAATGCTTCGCTGG 663
DB 191 HisLeuAspGlnValAlaAlaLeuHisTrpValGlnAspAsnIleAspAsnPheGly 210
QY 664 GACCCACACTGTGACCATGCTTTGGGAGTCCGCGGAGGCATAGTGTCTGCTGCT 723
DB 211 AspProGlySerValThrIlePheGlyGlnSerAlaGlyGlyLysValSerValLeu 230
QY 724 ATACTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGCTGAGC 783
DB 231 ValLeuSerProLeuAlaIaPheAsnLeuPheHisLysAlaIleSerGluSerGlyValAla 250
QY 784 ATCATCCTTACCTG---GAGGCCATATTAATGAAGAGTGAAGACCTGCAAGTGT 840
DB 251 LeuThrAlaGlyLeuValIaLysAsnThrArgProLeuAlaGluLysIleAlaValAl 270
QY 841 GCACATTTCTGTGATGAATGCGTCAAGCTGTGAGCGCTGAGGCTGAGGACAGCA 900
DB 271 Ser-----GlyCysLysSerThrThrSerAlaSerMetValHisCysLeuArgGln 287
QY 901 AAACCTTCAAGAGAGCTGTGACCTCAGCCAGCAAGAAACAAAGTCTTTCAT----- 951
DB 288 LysThrGlnGlnGluLeuLeuGlnThrThrLeuLysLeuAsnLeuPheSerLeuAspLeu 307
QY 952 -----CGAGTGTGATGTGCTTCTTT 975
DB 308 HisGlyAspSerArgGlnSerTyProPheValProThrValLeuAspGlyValValLeu 327
QY 976 CCTAATGACCTCTAGATCTATTGCTCAGAAAGCATTAAGCAATTCCTTCATCATC 1035
DB 328 ProLysMetProGlnGluIleLeuAlaGluLysAspPheAsnThrValProTyTrIleVal 347
QY 1036 GAGATCAATTAACCAAGAGTGTGCTTCTGCTGCT-----ATGAAGAGAGGCTCTGAGATC 1092
DB 348 GlyIleAsnLysGlnGluPheGlyTrpIleLeuProThrMetLeuAsnTyProProSer 367
QY 1093 CTCAGTGGCTCCAAACAGTCCCTGCTCCTCATCTGATTAACAAC-----ATCTG 1143
DB 368 AspMetLysLeuAspProMetThrAlaThrSerLeuLysSerSerPheLeuLeu 387
QY 1144 CACATCCCGCTCAGTATTTCACCTTGTGGTAAATGAATCTTCACAGACAGCAAGCTCC 1203
DB 388 AsnLeuProGlnAlaIleProValAlaValGluLysTyLeuAspHisThrAspAsp 407
QY 1204 CTGACTGAATCCGAGACAGTCTTCTGACTGTGCTTGGAGATGTGTTGTGTGCTCT 1263
DB 408 ProAspArgAsnLysAspGlnLeuLeuGlnLeuIleGlyAspValIlePheGlyValPro 427
QY 1264 GCATGATCAACAGCTGATATACAGAGATGCTGTGACCTGTCTACTTCTATGAGTTT 1323
DB 428 SerValIleValSerArgGlyHisArgAspAlaGlyAlaArgThrTyMetTyGluPhe 447

```

QY 1324 CGGACCGGCTCAGTGTCTTGAAGACAGAGCGGCTTTTGTCAAGCCGACGCT 1383
Db 448 GlnTyrArgProSerPheSerLeuMetLysProSerThrValAlaGlnPheGly 467
QY 1384 GATGAAGTCGGCTTGTGTGGTGGTGGCTTCTGAAGGGGACATTTGTTGGA 1443
Db 468 AaRgGlnIleTyrSerValPheGlnAlaProIleLeuAaRgLy----- 481
QY 1444 GGAGCCACGGAGAGAGAAATTACTGAGCCGGAAGATGATGAATACTGGCTACTTT 1503
Db 482 GlnTyrSerLysGlnIleValLeuSerLysMetMetLysPheThrAlaPhe 501
QY 1504 GCTCGAACCGGGAATCTTAATGGAGACGACTGTCTGTGGCGAGCTTAATCTGACT 1563
Db 502 AlaAaGlnGlyAaPheProSerGlnGlnGlnLeuProHISITrProGlnTyrAaRgGlnLys 521
QY 1564 GAGCAGTACTTCAGCTGACTTGAACATGAGCTTCGACAGACATCAAGACCGCG 1623
Db 522 GlnGlnTyrLeuGlnIleGlyAlaThrThrGlnGlnAlaGlnIleLeuLysGlnLysGln 541
QY 1624 GTGAGTTTTTGGACGACGACCATCCCGGATCCGTGCTGCTCCGACATGCTCCACAT 1683
Db 542 ValAlaPheTrpSerGln-----LeuValAlaMetLysPheLeuHISAla 556
RESULT 15
150367
carboxylesterase (EC 3.1.1.1) ES-10 precursor, microsomal - rat
N.Alternate names: hydrolase A
C.Species: Rattus norvegicus (Norway rat)
C.Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text_change 09-Jul-2004
C.Accession: S10367; S12468; S51202; S23460; S14361
R.Robbl, M.; Beaufay, H.; Octave, J.N.
Biochem. J. 269, 451-458, 1990
A>Title: Nucleotide sequence of cDNA coding for rat liver PI 6.1 esterase (ES-10), a car
A.Reference number: S10367; MUID: 90351366; PMID: 2386485
A.Accession: S10367
A:Molecule type: mRNA
A:Residues: 1-565 <ROB1>
A:Cross-references: UNIPROT:O9R135; EMBL:X51974
A>Note: 168-Gln, 247-Lys, 423-Met, and 506-Aa were also found
A>Note: the sequence from Fig. 4 is inconsistent with that from Fig. 5 in having 265-Lys
R.Robbl, M.
submitted to the EMBL Data Library, February 1990
A:Reference number: S12468
A.Accession: S12468
A:Molecule type: mRNA
A:Residues: 1-264, 'K', 266-565 <ROB2>
A:Cross-references: EMBL:X51974; NID:956898; PIDN:CAA36236.1; PID:956899
R.Morgan, B.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A>Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID: 7986098
A.Accession: S51202
A:Molecule type: Protein
A:Residues: 19-48 <MOR>
R.Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A>Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A:Reference number: S23460; MUID:92299008; PMID: 1606962
A.Accession: S23460
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-185, 'Q', 187-422, 'M', 424-505, 'N', 507-565 <MED>
A:Cross-references: EMBL:X65296; NID:957553; PIDN:CAA46391.1; PID:957554
R.Gaustad, R.; Sleiten, K.; Lowhaug, D.; Fornum, F.
Biochem. J. 274, 693-697, 1991
A>Title: Purification and characterization of carboxylesterases from rat lung.
A:Reference number: S14361; MUID:91190080; PMID: 2012599
A.Accession: S14361
A:Molecule type: Protein
A:Residues: 19-26, 'D', 28-37 <GAU>
C:Superfamily: cholinesterase; cholinesterase homology

C.Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-565/Product: carboxylesterase #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CH>
F:79,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:221,466/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,05e-78 Length: 565
Score: 1121.50 Matches: 245
Percent Similarity: 56.06% Conservative: 79
Best Local Similarity: 42.39% Mismatches: 199
Query Match: 35.38% Indels: 55
DB: 2 Gaps: 11

US-10-023-515-3 (1-1746) x S10367 (1-565)

QY 34 TGTGCTTTTCTGATTTCTCAGCCCTGTGGGACACAGACATGGGGAATACTGG 93
Db 8 TrpLeuPhe-----LeuAlaIaCythrAlaTrpGlyTyr----- 19
QY 94 CTTTGTGTAAGGCCACAGAGAACACGAGCTGGATGATTCAGGGCAAGTC 153
Db 20 ProSerSerPro---ProValAlaSerThrValLysGlyValLeuGlyLysTyrVal 38
QY 154 ACTGAGCTGGAGAGCCCTGCTGTAACGCTGTCCTCGGAGTCCCTTGTGCTGCC 213
Db 39 AsnIleuGlnGlyPheAlaGlnProValAlaAlaPheLeuGlyLysPhePheAlaLysPro 58
QY 214 CCGTGTGATCCCTGCAATTCAGAACCCGAGCTGCATGCCCTGGAGATTCAGTGA 273
Db 59 ProLeuGlySerLeuAaRgPheAlaProProGlnProAlaGlnProTrpAaPheValLys 78
QY 274 GAAGCCACTCTTACCTTAATTGTGCTCCAGAAC-----TCA 312
Db 79 AsnThrThrSerTyrProPrometCysSerGlnAaPheAlaValGlyGlnValLeuSer 98
QY 313 GAGTGGCTGCTTAAATCAACATGCTCAAGGTGATTAACCGGAATTCGAGTCA 372
Db 99 GlnLeuPheThrAsnAaGlyGlnAsnIleProLeuGlnPhe-----Ser 113
QY 373 GAAGATGCTCTTACCTGAACATCTATGCGCTGCCAGCCGATACAGGCTCCAGCTC 432
Db 114 GlnAaPheLysLeuTyrLeuAaPheValTyrThrProAlaSerPheThrLysAaSerArgLeu 133
QY 433 CCGCTCTTGTGTGTTCCAGAGAGTCCCTTCAAGCTGCTCAGCTCCATCTTGAT 492
Db 134 ProValMetValTrpIleAsnGlyGlyLeuValValGlyAlaSerThrTyrAsp 153
QY 493 GGGTCCGCGCTGGCTGCTTAAAGAGCATGCTGTTGTCGTCGTCAGTCCGCTAGGA 552
Db 154 GlnGlnValLeuSerAlaAsnGlnAaPheValValValThrIleGlnTyrAaGlyLeu 173
QY 553 ATATTGTGTTTTCACACATGGGATCAGCATGCTCCGGGGAATCGGGCTTCAGAGAC 612
Db 174 IleTrpGlyPhePheSerThrGlyAaRgGlnHISerArgLysIleAsnTrpGlnHISLeuAsp 193
QY 613 CAGGTGCTGCTGTCTGCTTGGGTCGCAAAACATGAGTCTTGGTGGGACCCGAGC 672
Db 194 GlnAlaAlaIleAsnHISTrpValGlnAaPheAsnIleAlaAsnPheGlyLysAaPheGly 213
QY 673 TCTGTGACATCTTTGGCGAGTCCGCGGAGCCATGAATGTTCTTAATCATGTCT 732
Db 214 SerValThrIlePheGlnGlyIleSerAlaGlyLysPheSerValSerAlaLeuValLeuSer 233
QY 733 CCATGAGCCAAAGGCTTATTCACAAAGCCATGAGAGAGTGGGTGGCCATCATCCCT 792
Db 234 ProLeuAlaLysAaPheAsnLeuPheHISArgAlaIleSerGlnSerGlyValValLeuThrSer 253
QY 793 TACTGAGGCCCATATTAATGAGAAAGAGAGAGACTCGAGGTGTTGACATTTCTGT 852
Db 254 AlaLeuIleThrThrAaPheLysProIleAlaAsnLeu-----IleAlaThrLeuSer 271

```
OY 853 GGTAAACAATGCGTCAGACTGTGAGCGCCCTGAGTGCCTGAGACAAACCTTCAAG 912
    |||
    |||
    |||
Db 272 GlyCysIsthrThrThrSerAlaValMetValHisCysLeuArgGlnIsthrGluAsp 291
    |||
    |||
    |||
OY 913 GAGCTGCTGACCCCTGAGCCAGAAAACAAAGCTTTC----- 948
    |||
    |||
    |||
Db 292 GluLeuGluIsthrSerLeuIsthrLeuAsnLeuPheIsthrLeuAspLeuGluIsthrAsnPro 311
    |||
    |||
    |||
OY 949 -----ACTGAGTGTGATGCGCTTCTTCTTCTAATGAGCCT 987
    |||
    |||
    |||
Db 312 LysGluSerTyrProPheLeuProThrValIleAspGlyValIleuProLysThrPro 331
    |||
    |||
    |||
OY 988 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATCGAGTCAATAAC 1047
    |||
    |||
    |||
Db 332 GluGluIleLeuAlaGlnIsthrSerPheAsnThrValProTyrIleValGlyIsthrLeuIsthr 351
    |||
    |||
    |||
OY 1048 CACGAGTGTGCTTCCTGCTGCTATGAAAGAGGCTCCTGAGATCCTCAGTGGC----- 1101
    |||
    |||
    |||
Db 352 GlnGluPheGlyTyrIleIsthrProThrLeuMetGlyTyrProLeuSerGlnGlyIsthrLeu 371
    |||
    |||
    |||
OY 1102 -----TCCAACAAGTCCCTTGCCCTCCATCTGATACAAAACATCCTGCACATCCCG 1152
    |||
    |||
    |||
Db 372 AspGlnIsthrAlaIsthrSerLeu---LeuTyrLysSerTyrProThrLeuIsthrSer 390
    |||
    |||
    |||
OY 1153 CCTCAGTATTGTGACCTGTGCTGATGATTAATCTTCATGACACAGACATCCTGACTGAA 1212
    |||
    |||
    |||
Db 391 GluIsthrMetIleProValValIleGlnIsthrTyrPheGlyIsthrAspAspProAlaIsthr 410
    |||
    |||
    |||
OY 1213 ATCCGAGACAGTCTTCTGAGCTTGTGAGATGTGTTCTTGTGGTCCCTGCACTGATC 1272
    |||
    |||
    |||
Db 411 ArgLysAspLeuPheGlnAspLeuValIleAspValIlePheGlyValProSerValMet 430
    |||
    |||
    |||
OY 1273 ACAGCTCGATATCAGAGATGCTGTGCACTGTCTACTTCTATGAGTTTCGGCACCGG 1332
    |||
    |||
    |||
Db 431 ValSerArgSerHisArgAspAlaGlyAlaProThrPheMetTyrGlnPheGluTyrArg 450
    |||
    |||
    |||
OY 1333 CCTGAGTCTTGAAGACGCGGCTTTGTCAAGCCGACGACGACGCTGATGAGTGC 1392
    |||
    |||
    |||
Db 451 ProSerPheValSerAlaMetArgProLysThrValIleGlyAspHisGlyAspGluLeu 470
    |||
    |||
    |||
OY 1393 CGCTTGTGTTCGGTGTGCTTCTGTAAGGGGACATTGTTATGTTTCGAAAGAGCCACG 1452
    |||
    |||
    |||
Db 471 PheSerValPheGlySerProPheLeuIsthr-----AspGlyAlaSer 484
    |||
    |||
    |||
OY 1453 GAGGAGGAGAACTTACTGAGCCGGAAGATGAAATACTGGGCTACTTGTCTGGAAC 1512
    |||
    |||
    |||
Db 485 GluGlnGluIsthrAsnLeuSerIsthrMetValMetLysTyrTyrAlaAsnPheAlaArgAsn 504
    |||
    |||
    |||
OY 1513 GGGATCTTAATGGAAGACCTGTCTCTGTGGCCAGCTTAATCTGACTGAGCAGTAC 1572
    |||
    |||
    |||
Db 505 GlySerProAsnGlyGlyLeuProHisTyrProGluTyrAspGlnLysGlnGlyTyr 524
    |||
    |||
    |||
OY 1573 CTCCAGCTGACTTAACAATGAGCCTCGGACAGACAGACTCAAGAAACCGCGGTGATTTT 1632
    |||
    |||
    |||
Db 525 LeuIsthrIleGlyAlaSerThrGlnAlaIleGlnArgLeuIsthrAspLysGlnValAlaPhe 544
    |||
    |||
    |||
OY 1633 TGGACCAAGACCATCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTCT 1686
    |||
    |||
    |||
Db 545 TrpSerGlu-----LeuArgAlaLysGlnAlaIleGlnGluPro 557
    |||
    |||
    |||
```

Search completed: June 15, 2005, 08:55:41
Job time : 60.53 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 02:39:48 ; Search time 155.861 Seconds
(without alignments)
11472.925 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 3170
Sequence: 1 atgcacacagagcactactctc.....ttctcttctgtctctctga 1746

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 segs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+np.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/6664091/rxnat.14062005.133459.9335/app_query.fasta.1.4238
-DB=UniProt_03 -OPMT=fastaan -SUPFIX=n2p.rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEADSIZE=500 -MITLEN=0 -MAYLEN=2000000000
-USER=6664091@cgn2.1.389 @rxnat.14062005.133459.9335 -NCPU=6 -ICPU=3
-NO WMAP -LANGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONOLIOG
-DEV TIMEOUT=120 -WAPN TIMEOUT=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03 :
1: uniprot_sprotc :
2: uniprot_trembl :*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2915	92.0	575	2	Q6NT32
2	2606	82.2	525	2	Q96DN9
3	2454	77.4	469	2	Q8N8C8
4	2250	71.0	575	2	Q6AW47
5	2094.5	66.1	575	2	Q6AW46
6	1985	62.6	545	2	Q81034
7	1676.5	52.9	361	2	Q95KH3
8	1298.5	41.0	557	1	SASB ANAPL
9	1282.5	40.5	561	2	Q91WGO
10	1281	40.4	559	2	Q8BK48
11	1276.5	40.3	559	2	Q35533
12	1271	40.1	559	2	Q8BW97
13	1266.5	39.9	556	2	Q6PDB7
14	1263.5	39.8	561	2	Q6GM54
15	1262.5	39.8	561	2	Q70177
16	1260.5	39.8	560	2	Q8R097

17	1257.5	39.7	554	2	Q35535	035535	rattus norv
18	1239.5	39.1	552	2	Q6P306	06P306	xenopus tro
19	1238	39.1	568	2	Q8K3R0	08K3R0	rattus norv
20	1232.5	38.9	558	2	Q8OZK3	08OZK3	mus musculu
21	1231.5	38.8	532	1	EST2_RABIT	P14943	oryctolagus
22	1230	38.8	559	1	EST2_HUMAN	U00748	homo sapien
23	1213.5	38.3	561	2	Q70631	070631	rattus norv
24	1213.5	38.3	566	2	Q766D7	0766D7	felis silve
25	1213.5	38.3	566	2	Q864S9	0864S9	felis silve
26	1210.5	38.2	557	2	Q640T6	0640T6	xenopus tro
27	1159	37.8	561	1	EST1_MESAU	Q64419	mesocricetu
28	1191	37.6	586	2	Q61PK9	061PK9	homo sapien
29	1187.5	37.5	568	2	Q61PK9	061PK9	mus musculu
30	1187.5	37.5	572	2	Q810S9	0810S9	mus musculu
31	1186.5	37.4	554	1	ESTM_MOUSE	Q95N05	canis fami
32	1183	37.3	565	2	Q95N05	Q95N05	canis fami
33	1179.5	37.2	571	2	Q6UW88	06UW88	homo sapien
34	1179	37.2	568	2	Q726J1	0726J1	homo sapien
35	1175	37.1	565	2	Q81DZ9	081DZ9	homo sapien
36	1171.5	37.0	568	2	Q8VCU1	08VCU1	mus musculu
37	1169.5	36.9	561	2	Q91WU0	091WU0	mus musculu
38	1164	36.7	565	2	Q97582	097582	sus scrofa
39	1162.5	36.7	555	2	Q6GMU1	06GMU1	brachydanio
40	1162.5	36.7	566	1	EST1_PIG	Q29550	sus scrofa
41	1159.5	36.6	566	2	Q464Z1	Q464Z1	macaca fasc
42	1150.5	36.3	565	1	EST1_RABIT	P12373	oryctolagus
43	1150	36.3	554	1	ESTM_MOUSE	P23953	mus musculu
44	1149	36.2	565	2	Q35534	Q35534	mesocricetu
45	1144	36.1	549	1	EST1_RAT	P10959	rattus norv

ALIGNMENTS

RESULT 1
Q6NT32 PRELIMINARY; PRT; 575 AA.
AC Q6NT32;
ID Q6NT32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE FLJ1547 protein.
GN Name=FLJ1547;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman A.C., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;

RA Director MGC Project;
RU Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: BC069501, AAH69501.1; -.
DR HSBP: F12337; 1K4Y.
DR GO: GO:0003824; F: catalytic activity; IEA.
DR InterPro: IPR002018; CarboxesteraseB.
DR InterPro: IPR000379; Ser esters.
DR Pfam: PF00135; Coesterase_1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE. 575 AA; 63926 MW; 6F5B735BDFEC9C09 CRC64;

Alignment Scores:
Pred. No.: 2,25e-208 Length: 575
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservative: 3
Best Local Similarity: 97.02% Mismatches: 6
Query Match: 91.96% Indels: 8
DB: 2 Gaps: 1

US-10-023-515-3 (1-1746) x Q6NTJ2 (1-575)
QY 34 TGGTCTTTTCTGATTCCTCCAGCCCTGTGGGACACAGACAGTGGGAAAACTGGG 93
DB 14 TrrlalaletrpValleuhalalaPro-----Thrlysgly 25
QY 94 CTTTCTGCGAAGGGGCGACAGAGAAACACAGGCTGGGATGGATTGAGGGCAAGCAAGTC 153
DB 26 ProSerlaagluglYrproglnaArgantrArgleuglYtrp1le6lndglYybsglInval 45
QY 154 ACTGTGCTGGGAAGCCCTGCGCTGTGGAACGTGTTCTCGAGTCCCGCTTGTGCTGCC 213
DB 46 ThrValleuGlYSerProValProValasnValPheleuGlYValProPhehalalaPro 65
QY 214 CCGGTGGAGATCCCTGAGATTAGCAACCCGACGCTGCATGCGCCCTGGAGTAATTGCGA 273
DB 66 ProleuGlYSerleuArgPheThrAnPProglInProAlaSerProtrpAsnleuArg 85
QY 274 GAAGCCACTCTCAACATTTTGGCTCCGAACTCGAGATGCTGCTTTAATCAA 333
DB 86 GluAlaTrpSerYrProAsnleuCyseuGlInAsnSerGlYtrpLeuLeuAspGln 105
QY 334 CACATGCTCAAGGTCATTAACCCGAAATTCGAGTGTCAAGAGACTGCTTACCCGAAC 393
DB 106 HisMetleuYrValHisYrProlysrPheGlYValSerGlAspCyseuLeuYrleuAsn 125
QY 394 ATCTATGCGCTGCCACGCGCGATACAGGCTTCAAGCTCCCGTCTTGAGTGTGCCA 453
DB 126 IleYrAlaProAlaHisAlaPheThrGlYSerlyseuProValleuValTrpPhePro 145
QY 454 GGAGGTGCTTCAAGATGGCTCAGCTCCATCTTTGATGGTCCGCCCTGGCTCCAT 513
DB 146 GlYglYAlaPheleuYrThrGlYSerAlaSerIlePheAspGlYSerAlaLeuAlaYr 165
QY 514 GAGGACTGTGCTGTGCTGCTCCAGTACCGGCTAGAGAAATTTGGTTTCTTACACACA 573
DB 166 GluAspValleuValValValGlnYrArgleuGlYIlePheGlYPhePheThrThr 185
QY 574 TGGGATCGCATGCTCCCGGGAATCGGCTTCAAGAACAGAGTGGCTGTCTGCTGG 633
DB 186 TrpAspGlnHisAlaProGlYAsnTrpAlaPheleuYrAspGlnValAlaAlaLeuSerTrp 205
QY 634 GTCCAGAAAGAACATGAGATTCTGGTGGGAGCCCACTCTGAGACATCTTTGGCGAG 693
DB 206 ValGlnYrAsnIleGlnPhePheGlYglYAspProSerSerValThrIlePheGlYglu 225
QY 694 TCCGCGGAGCCATAAGTGTCTTAGTCTTACTGTCTCCCATGCGCAAGCTTATTC 753
DB 226 SerAlaGlYAlaIleSerValSerSerleuIleuSerProMetAlaYlYglYleuPhe 245
QY 754 CACAAAGCATCATGAGAGTGGGGTGGCCATCATCTCTTACTGAGAGCCCATGATTTAT 813

DB 246 HisYrAlaIleMetleuSerGlYAlaAlaIleIleProYrleuGlInAlaHisPheYr 265
QY 814 GAGAAAGTAGAGACCTCGAGGTGGTGGACATTTCTGTGTAAAGATCGTCAGACTCT 873
DB 266 GluYrSerGlAspPheuGlInValAlaHisPheCyseuGlYAsnAsnAlaSerAspSer 285
QY 874 GAGGCGCTCTGAGGTGCTTGAAGACAAACCTTCAGAGACTGTGACCTCAGCCAG 933
DB 286 GluAlaLeuLeuArgCyseuArgThrYrProSerlyseuLeuLeuThrleuSerGln 305
QY 934 AAAACAAAGCTTTTCACTCGAGGTGGTGGTGGCTTTCTTCTTAAGGCTCTAGAT 993
DB 306 YrThrlysrSerPheThrArgValAlaAspGlYAlaPhePheProAsnGlnProleuAsp 325
QY 994 CTATTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACGAG 1053
DB 326 LeuLeuSerGlnYrAlaPheYrAlaIleProSerIleIleGlYValAsnAsnHisglu 345
QY 1054 TGTGGCTTCCGCTCCCTATGAGAGAGGCTCTGAGATCTTGAAGGCTCCCAAGATCC 1113
DB 346 CyseuYrPheleuLeuProMetlyseuAlaProGlnIleleuSerGlYSerAsnlysr 365
QY 1114 CTTGGCTTCCATCTGATACAAGACATCTCGACATCCGCGCTCAGTATTTGGACCTTGTG 1173
DB 366 LeuAlaLeuHisleuIleGlnAsnIleleuHisIleProProGlnYrleuHisAlaVal 385
QY 1174 GCTAATGAATCTTCATGACACAGACTCCCTGACTGAATTCGAGACAGTCTTGAC 1233
DB 386 AlaAsnGluYrPheHisAspYrHisSerleuThrGluAlaArgSerleuLeuAsp 405
QY 1234 TTTGCTTGAGATGAGTTTGTGGTGGCTCGACATGATCAGAGCTCGATACAGAGAT 1293
DB 406 LeuLeuGlYAspValPhePheValAlaProAlaLeuIleTrpAlaArgYrHisIaArgAsp 425
QY 1294 GCTGTGACACTGTCTACTTTATGATTTGCGGACCGGCTCAGTGGCTTTGAAGACAG 1353
DB 426 AlaGlYAlaProValYrPheYrGlYsrPheArgHisAspProGlnCyseuPheGluAspThr 445
QY 1354 AAGCGGCTTTTTCGAAAGCCGACACAGCTGATGAAGTCCGCTTGTGTGGTGGTCC 1413
DB 446 LysProAlaPheValYrValAlaAspHisAlaAspGlnValAlaArgPheValPheGlYglYAla 465
QY 1414 TTCTGAAAGGGGACATGTTATGTTGAGAGAGCGACGAGGAGGAGATTTCTGAGC 1473
DB 466 PheleuYrsglYAspIleValaMetPheGlYAlaTrpGluGluGluYrleuLeuSer 485
QY 1474 CGAAGATGATGAATACTGGGCTAATCTTGTCCGAACCGGGAATCTTAATGGAAACGAC 1533
DB 486 ArgYsrMetleuCyYrYrTrpAlaThrPheAlaArgThrGlYAsnProAsnGluYAsnAsp 505
QY 1534 CTGTCTCTGTGGCGACGTTAATATCTGATGACAGAGTACCTTCAGCTGGAATTAACATG 1593
DB 506 LeuSerleuTrpProAlaYrAsnleuThrGlYndglYrleuGlnleuAspLeuAsnMet 525
QY 1594 AGCTTCGACAGAGACTCAAGAACCGGAGTGAATTTTGGACGAGACATCCCTGTG 1653
DB 526 SerleuGlYglInArgleuYrsglYrProArgValaAspPheTrpTrpSerThrIleProleu 545
QY 1654 ATCTGTCTGCTCCGACATGCTCCACAGTCTCTTTCTTCTTAATCTTCTCTCTCT 1713
DB 546 IleleuSerAlaSerAspMetleuHisSerProleuSerSerleuThrPheleuSerleu 565
QY 1714 CTCAGGCTTCTTTCTTTTCTTTGGCTCCT 1743
DB 566 LeuGlnProPhePhePheCyseuAlaPro 575

RESULT 2
O96DN9 PRELIMINARY; PRT; 525 AA.
AC O96DN9; (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)

DE Hypothetical protein FLJ31547.

GN Name=FLJ31547;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RP PubMed=14702039; DOI=10.1038/ng1285;

RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,

RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,

RA Setine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishi S.,

RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,

RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,

RA Sudo H., Hosioki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,

RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,

RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,

RA Ninomiya K., Iehabashi T., Yamashita H., Murakawa K., Fujimori K.,

RA Tanai H., Kimura M., Watanabe M., Hirooka S., Chida Y., Ishida S.,

RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T., Kusano J.,

RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,

RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,

RA Masashiro K., Yuki H., Oshima A., Sasaki N., Aotaka S.,

RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,

RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,

RA Nakagawa S., Sanoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,

RA Hishiki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,

RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,

RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujizara T.,

RA Ono T., Yamada K., Fujii Y., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,

RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,

RA Togoashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,

RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,

RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;

RT "Complete sequencing and characterization of 21,243 full-length human

RT cDNAs.";

RL Nat. Genet. 36:40-45(2004).

RN [2]

RP SEQUENCE FROM N.A.

RP TISSUE=PCR rescued clones;

RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RX Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Sheman C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA DiCicco L., Marusina K., Farmer A.J., Rubin G.M., Hong L.,

RA Steplenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Teshlinski S., Cantini P., Prange C.,

RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,

RA Bosak S.A., McSwan P.J., McKeen K.J., Malek J.A., Gamarate P.H.,

RA Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Ketterman M., Madan A.C., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez M.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Kravitzneski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.B.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL [3]

RP SEQUENCE FROM N.A.

RP TISSUE=PCR rescued clones;

RA Strausberg R.;

RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.

OR EMBL; AK056109; BAB71094.1; -

OR EMBL; BC069548; AA69548.1; -

DR	HSSP: P12337; 1k4Y.		
DR	GO: GO:0016787; F:hydrolyase activity; IEA.		
DR	InterPro: IPR002018; Carbohydraseb.		
DR	InterPro: IPR000379; Ser_castrs.		
DR	Pfam: PF00135; Coesterase; 1.		
DR	PROSITE: PS00122; CARBOXYLSTEPEASE_B_1; 1.		
DR	PROSITE: PS00941; CARBOXYLSTEPEASE_B_2; 1.		
KM	Hydrolase; Hypothetical protein.		
SQ	SEQUENCE 525 AA; 58200 MW; 7724878B8A22F215 CRC64;		
 Alignment Scores:			
Pred. No.:	2,31e-185	Length:	525
Score:	2606.00	Matches:	503
Percent Similarity:	88.77%	Conservative:	3
Best Local Similarity:	88.25%	Mismatches:	6
Query Match:	82.21%	Indels:	58
DB:	2	Gaps:	2
 US-10-023-515-3 (1-1746) x Q96DN9 (1-525)			
OY	TGSGCTTTTCCGTGAATTCAGGCCCTGTTGGACACAGACAGTGCGGAAAACTGGG	93	
Db	14 Trpalailetrpvalileualaialapro-----Thrlsgly	25	
OY	CCTTGTGTGAAGGGCCACAGAGAAACCAGGCTGGATTCAGGGCAGAAGTC	153	
Db	26 Proserialagluglprproglnatrgasnthraagleuglyttrplldnglylsghnal	45	
OY	ACTGAGCTGGGAAGCCGTGTGCCTGTGAACGATTCCTGSAGTCCCCTTGTGCTCCC	213	
Db	46 Thrvalleuglyserprovalprovalaenvalphenelglyalprotheaalaaipro	65	
OY	CCGCTGGGATCCCTCGATTNACGAACCCGACGCTGTACCGCCCTGGATAACTGGCA	273	
Db	66 Proleuglyserleudargphethrhanproglinproalaaserprotrphaenleuarx	85	
OY	214 GAAGCACCTCTACCCCTTAATTTGGTCCCTCCAGAACSTAGAGTGGGTCTTAGATCAA	333	
Db	86 Glualatrhsertryproasnleucyuleuglnaanserglutripheuleueaergln	105	
OY	CACATGCTCAAGGTGCATTAACCCGAAATTCGGAGTGCAGAAAGCTGCTACTGTAAC	393	
Db	106 Hlwecleulyvalnhisttryprolyubneeglyalsegclunbpcyleutyryleuaen	125	
OY	ATCTATGGCGCTGCCACGCGCATACAGGCTCCAAGCTCCCGCTTGTGTGTTCCA	453	
Db	126 lletrytlarproaliahaiaahamrhnrglyserlyvleuovalleuualtriphepro	145	
OY	GGAGTGCTTCAAGACTGAGCTGACGCTTCATCTTTGATGGGTCCGCGCTGCTGCTAT	513	
Db	146 Glyglylaalpneluthrclyserlalaserllpheabpglyserllaaleualaaalyr	165	
OY	GAGGACGTGTGTTGTGTGTCAGTAACGGGCTAGAAATTTGGTTTTCTTACACACA	573	
Db	166 Gluarpvalleivalvaivaivaiglntryatrglenglyllephenolyphrephnthr	185	
OY	TGGATTCAGCATGCTCCGGGGAACTGGGCTTCAAGGACAGAGGTGCTGCTGTCTGG	633	
Db	186 Ttrpargelnmhalaproclyasnttrplaaphelyaabrglnvalaaalaleuserttr	205	
OY	GTCCAGAAAGAACATGAGTCTTTCGATGGGGAACCCAGACTGTGTGACCATCTTTGGGAG	693	
Db	206 Valglnlybaenllgluphenehgylgyaerprosserervalthrillepnehglyn	225	
OY	TCCGGGGAGCCATAAGTGTCTTAGTCTTAAAGTGTCCCATGTGGCCAAAGGCTATTTC	753	
Db	226 Seratalaglyalalleservalasertserleuileleneserprometrlaubygilylepne	245	
OY	CACAAAGCATCATGAGAGAGTGGGCTGGCATATCCCTTACCTGGAGGCCATGATTAT	813	
Db	246 Hislysalaitiemecgluserclyalaalaitelleprototytleundulaaniabaptyr	265	
OY	GAGAAAGATGAGAACCTTGAGGGTGGTGTGACATTTCTGTGTAAACAATGCGTCAGACTT	873	

Accession	Gene	Species	Protein	Length	Start	Stop	Score	E-value	Bit Score	Identities	Positives	Gaps	Conserved Domains	Annotations	Accession	Gene	Species	Protein	Length	Start	Stop	Score	E-value	Bit Score	Identities	Positives	Gaps	Conserved Domains	Annotations
D8	266	GIUYSSEGI	Leu	285	1	285	100	0.0	100	100	100	0	100	100	D8	266	GIUYSSEGI	Leu	285	1	285	100	0.0	100	100	100	0	100	100
QY	874	GAGCCCTG	Leu	933	1	933	100	0.0	100	100	100	0	100	100	QY	874	GAGCCCTG	Leu	933	1	933	100	0.0	100	100	100	0	100	100
Db	286	GIUATALE	Leu	305	1	305	100	0.0	100	100	100	0	100	100	Db	286	GIUATALE	Leu	305	1	305	100	0.0	100	100	100	0	100	100
QY	934	AAAAAAGT	Leu	993	1	993	100	0.0	100	100	100	0	100	100	QY	934	AAAAAAGT	Leu	993	1	993	100	0.0	100	100	100	0	100	100
Db	306	LYSTHLYS	Leu	325	1	325	100	0.0	100	100	100	0	100	100	Db	306	LYSTHLYS	Leu	325	1	325	100	0.0	100	100	100	0	100	100
QY	994	CTATTGTG	Leu	1055	1	1055	100	0.0	100	100	100	0	100	100	QY	994	CTATTGTG	Leu	1055	1	1055	100	0.0	100	100	100	0	100	100
Db	326	LEULEUSE	Leu	345	1	345	100	0.0	100	100	100	0	100	100	Db	326	LEULEUSE	Leu	345	1	345	100	0.0	100	100	100	0	100	100
QY	1054	TGTGGCTT	Leu	1113	1	1113	100	0.0	100	100	100	0	100	100	QY	1054	TGTGGCTT	Leu	1113	1	1113	100	0.0	100	100	100	0	100	100
Db	346	CYSGLYP	Leu	365	1	365	100	0.0	100	100	100	0	100	100	Db	346	CYSGLYP	Leu	365	1	365	100	0.0	100	100	100	0	100	100
QY	1114	CTTGCCCT	Leu	1173	1	1173	100	0.0	100	100	100	0	100	100	QY	1114	CTTGCCCT	Leu	1173	1	1173	100	0.0	100	100	100	0	100	100
Db	366	LEUATALE	Leu	385	1	385	100	0.0	100	100	100	0	100	100	Db	366	LEUATALE	Leu	385	1	385	100	0.0	100	100	100	0	100	100
QY	1174	GCTAATGA	Leu	1233	1	1233	100	0.0	100	100	100	0	100	100	QY	1174	GCTAATGA	Leu	1233	1	1233	100	0.0	100	100	100	0	100	100
Db	386	ALAAENG	Leu	405	1	405	100	0.0	100	100	100	0	100	100	Db	386	ALAAENG	Leu	405	1	405	100	0.0	100	100	100	0	100	100
QY	1234	TTGCTTGA	Leu	1293	1	1293	100	0.0	100	100	100	0	100	100	QY	1234	TTGCTTGA	Leu	1293	1	1293	100	0.0	100	100	100	0	100	100
Db	406	LEULEUG	Leu	424	1	424	100	0.0	100	100	100	0	100	100	Db	406	LEULEUG	Leu	424	1	424	100	0.0	100	100	100	0	100	100
QY	1294	GCTGGTGA	Leu	1353	1	1353	100	0.0	100	100	100	0	100	100	QY	1294	GCTGGTGA	Leu	1353	1	1353	100	0.0	100	100	100	0	100	100
Db	424	-----	Leu	424	1	424	100	0.0	100	100																			

OC Mammalia, Eutheria, Primates, Catarrhini, Homiinae, Homo.
 OX NCBI_taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsushi K., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai T., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yasuda T., Iwanaga T., Magatsuma M., Shiroi A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Niimura Y. K., Ishihashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano Y.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togliya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Muesshino K., Yuki F., Oshima H., Saeki N., Aotaka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Teraishi Y., Suzuki O.,
 RA Hishigaki S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki M., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Omori Y.,
 RA Okabeata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Katsurai R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Saeki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL, AK090997; BAC03565.1; -
 DR HSSP; P12377, 1K4Y.
 DR GO: GO:0016787; F-hydrolase activity; IEA.
 DR InterPro: IPR002018; CarboxesteraseB.
 DR InterPro: IPR000379; Ser_ests.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 KW Hydrolase.
 SQ SEQUENCE 469 AA; 52312 MW; 7A188F68C10A9080 CRC64;

 Alignment Scores:
 Pred. No.: 4,72e-174 Length: 469
 Score: 2454.00 Matches: 467
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.57% Mismatches: 0
 Query Match: 77.41% Indels: 0
 Gaps: 2

 US-10-023-515-3 (1-1746) x Q8NBCE (1-469)

 QY 337 ATGCTCAAGGTGATTAACCGGAATTCCGAGTGTCAAGAACTGCTTACTGTAACATC 3366
 Db 1 MetLeuLyValHisIerYProLybPheGlyValSerGIuaBpCybLeuTYrLeuAsnIle 20

 QY 397 TAGGCGCTGCCAAGCGGATACAGAGTCGCAAGCTCCCGGCTGGTGGTGGTCCCGAGA 4566
 Db 21 TyAlaPbAlaHisAlaAspHisIerYbSerYbLeuPProValLeuValTrrPheProGly 40

 QY 457 GGTGCTTCAAGACTGAGCTCAGCTTCAATCTTGTATGGGTCCGCGCTGAGTGGCTATGAG 5166
 Db 41 GlYAlaPheLybIerThcGlySerAlaSerIlePheAspGlySerAlaLeuAlaAlaIerGlu 60

 QY 517 GACGTGCTGGTGTGGTGTGCTCCAGTACGGGCTAGGAATATTGGTTTCTTACCACATGG 5766


```

Db      145 ProGlyAlaLpheGluThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 164
Qy      511 TATGAGAGAGCTGCTGTTGCTGCTCGACGACCGGCTAGAAATATTGGTTCTTCAAC 570
Db      165 TrrGluAerValLeuIleValThrThrGlnTyrArgLeuGlyIlePheGlyPhePhe 184
Qy      571 ACATGGATCAGCATGCTCCGGGGAACCTGGCTTCAAGACACGAGTGGCTGCTGTC 630
Db      185 ThrGlyAerGlnIleAlaProGlyAsnTrpAlaPheLeuAspGlnLeuAlaLeuThr 204
Qy      631 TGGGTCCAGAAAGACATCGAGTTCTTCGGTGGGAGACCCGAGCTGTGACATCTTTGGC 690
Db      205 TrpValGlnGluAsnIleGluPhePheGlyGlyAerProIleSerValThrIlePheGly 224
Qy      691 GAGTCCGGGAGCCGCTAAGTGTCTTACTCTTAATCTGTCTCCATGGCCAAAGCTTA 750
Db      225 GluSerAlaGlyAlaIleSerValSerGlyLeuValLeuSerProMetAlaSerGlyLeu 244
Qy      751 TTCCCAAGACCATGAGAGAGTGGGGGCGCATCATCCCTTACCTGGAGCCCATGAT 810
Db      245 PheHISlybAlaIleMetGluSerGlyValAlaIleIleProPheLeuArgAlaProAsp 264
Qy      811 TATGAAAGAGTGAAGACCTTGACGTTGGTGCATTTCTGTGTAAACAATCGTCAGAC 870
Db      265 AepGluArgAsnGluAerLeuGlnValIleAlaArgIleCyArgGlyCyAsnValSerAsp 284
Qy      871 TCTGAGGCGCTGCTGAGGTGCTGAGACAAACCTCCAGAGACTGTGACCTCGAC 930
Db      285 SerValAlaLeuLeuGlnCysLeuArgAlaIlySerSerGluGluLeuLeuAspIleAsn 304
Qy      931 CAGAAACAAAGTCTTTCATCTCGATGAGTGAAGTGGCTTCTTCTTAATGAGCTCTA 990
Db      305 LysIlyThrIlySerPheThrArgValAlaAspGlyPhePhePheProAspGluProLeu 324
Qy      991 GATCTATTGTCTCAGAAACATTTAAACAAATCTTCCTCATGAGTGAATCAATACAC 1050
Db      325 AepLeuLeuThrGlnIlyThrPheAsnSerIleProSerValIleGlyValAsnAsnHis 344
Qy      1051 GAGGTGGCTTCCGCTGCTGCTTGAAGAGAGCTCTGAGATCTCTAGTGGCTCCAAAG 1110
Db      345 GluCybGlyPheLeuLeuProMetLysGluPheProGluIleLeuGlyGlySerAsnLys 364
Qy      1111 TCCCTTGGCCCTCATCTGATACAACATCTGACATCCGCGCTCAGATATTGGACCTT 1170
Db      365 SerLeuAlaLeuHisLeuIleHisArgValLeuHisIleProAsnGlnIlyLeuTyrIleu 384
Qy      1171 GTGCTTAATGAATACCTTCATGACACAGACATCCCTGACTGAATCCAGACAGCTTTCTG 1230
Db      385 ValAlaAerGlnTyrPheTyrAsnLysHisSerProValGluIleArgAspSerPheLeu 404
Qy      1231 GACTTGTGGAGATGCTTTTGTGTGCTCCCTGACATGACATCAGCTGCATATCACGA 1290
Db      405 AepLeuLeuGlyAerValPhePheValValProGlyValIleThrAlaArgTyrHisArg 424
Qy      1291 GATGCTGATGACCTGTCTACTTCTATGAGTTTGGACACCGGCTCAGATGCTTGAAGAC 1350
Db      425 AspIleGlyAlaProValTyrPheTyrGluPheGlnHisProProGluCyLeuLysAsp 444
Qy      1351 ACGAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTGTGTGGTGTG 1410
Db      445 ThrArgProAlaPheValIlyAlaAspHisSerAspGluIleArgPheValPheGlyGly 464
Qy      1411 GCCTTCCGAAAGGGGAGCATTTATGTTGGAAGAGACAGGAGGAGGAAGATTTCTG 1470
Db      465 AlaPheLeuLysGlyAsnIleValAlaMetPheGlnGlyAlaTrpGluGluLysLeuLeu 484
Qy      1471 AGCGGAAGATGATGAATATCTGGGCTACTTGTGCTCGAACCGGGAATCTTAATGGGAAC 1530
Db      485 SerArgLysMetMetLysGlyTrrpAlaAsnPheAlaGlnThrGlyAerProAsnGlyGlu 504
Qy      1531 GACTGTCTCTGTGGCAGCTTATTAATCTGACGATGACGATCTCCAGCTGGACTTGAAC 1590

```

```

Db      505 GlyLeuProLeuTrpProAlaTyrSerGlnSerGlnGlnIlyrLeuLysLeuAspLeuAsn 524
Qy      1591 ATGAGCTTCGACAGAGACTTAAGAACCGCGGTGAGATTTTGGACACGACCATCCCC 1650
Db      525 IleSerValGlyGlnLysLeuLysGlnGlnIlyValGluPheTrpSerAspThrLeuPro 544
Qy      1651 CTGATCCGTCGCTCCGACAGATGTCACAGTCTCTTCTTCTTCACTTCTCTCT 1710
Db      545 LeuIleMetSerMetSerThrAlaProProGlyProProValProLeuLeuSerLeuSer 564
Qy      1711 CTCTTCAGCCCTTCTTTCTTTTGTGCTCTCT 1743
Db      565 ValLeuLeuProPheLeuPheSerSerAlaPro 575

```

RESULT 5

```

Q6AM46 PRELIMINARY; PRT; 575 AA.
AC Q6AM46;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxin;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBT_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyazaki M., Yamashita T., Taira H., Suzuki A.;
RT "cauxin family protein."
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AB16393; BAD5016.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR002018; CarboxylesteraseB.
DR InterPro; IPR000379; Ser esterase.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR HydroLase.
SQ SEQUENCE 575 AA; 6416 MW; 3EB9D85981D9DE0A CRC64;

```

Alignment Scores:

```

Pred. No.: 3,13e-147 Length: 575
Score: 2094.50 Matches: 390
Percent Similarity: 80.67% Conservative: 69
Best Local Similarity: 68.54% Mismatches: 99
Query Match: 66.07% Indels: 11
DB: 2 Gaps: 1

```

US-10-023-515-3 (1-1746) x Q6AM46 (1-575)

```

Qy      34 TGTGCTTTTTCCTGATTCACAGCCCTGTGGACACAGAGTGGGGAACAACTGGG 93
Db      17 TrpIlePheGlyAlaIleIleGlu-----Gly 25
Qy      94 CTTTCTGCTGAAGGCGCCACAGAGAACACAGCTGGGATGATTCAGGGCGAAGCAAGTC 153
Db      26 SerValThrIluGluProHisArgTyrThrLysLeuGlyTrpValGlnGlyLysAla 45
Qy      154 ACTGTGTGGGAAGCCCTGCGTGAACGTGTCTCGAGAGCCCTTGTGCTGCC 213
Db      46 ThrValLeuGlyArgLeuGluProValAsnValPheLeuGlyIleProPheAlaAlaPro 65
Qy      214 CCGCTGGATCCCTGCGATTACGAACCCGACGCTGCATCGCTGGGATTAATTGGCA 273
Db      66 ProLeuGlyProLeuArgPheSerLysProGlnProProIleProTrpAspAsnLeuArg 85
Qy      274 GAAGCCACCTCCCAACCTAATTGTGCTTCAGAACTCAGAGTGGCTGCTTAAAGATCAA 333
Db      86 GluAlaThrAlaTyrProAsnLeuCybPheGlnAsnLeuGluTrpPhePheIleTyrGln 105

```

QY	333	TACATGCTCAAGGTCATTACCCGAAATTTGGAGTGTCAAGAACTGCTCTACTGTGAC	393
	106	AsnLeuLeuLysValSerTyrPro1LeuGlnMetSerGlnAspCysLeuTyrLeuAsn	125
QY	394	ATCATATGGCTGGTGGCCACGCGGATACAGAGCTTCAGAGCTCCCGCTGGTGGTCCCA	453
	126	IleTyrHisProCysHisAlaAsnAsnGlySerSerLeuProValMetValTyrIlePro	145
QY	454	GGAGGTGCTTCMAAGTGGCTGACCTTCATCTTTGATGGGTCCGCTGGCTGTAT	513
	146	GlyGlyGlyPheGlnThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaValTyr	165
QY	514	GAGGACGTCGTGTGTGGTGGTCCAGTACCGCGGCTAGGAATATTTGGTTTTCACCA	573
	166	GlnAspValLeuValValThrIleGlnTyrGlnGlyIlePheGlyPhePheThrThr	185
QY	574	TGGATATCAGATGCTCCGGGAGACTGGGCTTCAGAGACAGATGGCTGTGGTCTGG	633
	186	GlnAsnGlnHisAlaProGlyAsnThrAlaPheThrPheArgGlnLeuAlaIleLeuThr	205
QY	634	GTCCAGAAAGATCGAGTTCTTGGTGGGAGCCCGACCTGTGTACCATTTTGGCGAG	693
	206	ValArgGlnAsnIleLeuTyrPheGlyGlyAsnProAspSerValThrIlePheGlyAsn	225
QY	694	TCCCGCGGAGCGCATAGCTTTCTTAGTCTTATCTGTCTCCCAATGGCCAAAGCTTATC	753
	226	SerAlaGlyAlaIleSerIleSerSerLeuLeuSerProLeuSerAlaAspLeuPhe	245
QY	754	CACAAGGTCATCAGAGAGTGGGGTGGCCATCATCTTACCTGAGAGGCCCATGATAT	813
	246	HisArgAlaIleMetGlnSerGlyAlaAlaIleIleProSerLeuLysSerSerAsn	265
QY	814	GAGAAAGTGAAGACTGCGACAGTGGTGGACATTTCTGTGGTAAACATGCTGACACTT	873
	266	AspLeuLysHisAspLeuGlnValValAlaAsnValCysAspCysAsnValSerAspSer	285
QY	874	GAGGCTGCTGAGAGTGGCTGAGACAACCTTCAGAGAGCTGTGACCTCAGCCAG	933
	286	LysAlaLeuLeuLysCysValLeuAlaGlyLysSerSerLeuGlnLeuMetSerLeuSerGln	305
QY	934	AAACCAAGTCTTTCACCTCGAGTGGTGTATGGTGTCTTCTCCAAATGAGCTCTGAT	993
	306	LysAlaLysSerPheThrAlaValAlaAspIleSerPhePheSerGlnGlyProLeuGln	325
QY	994	CTATTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCCATCAGACGAGTCAATTAACACAG	1053
	326	LeuLeuSerGlnLysThrLeuLysIleValProSerIleIleGlyValaAsnAsnGlnGln	345
QY	1054	TGTGGCTTCTGCTGCTATGAGAAGAGCTCTGAGATCTTCAGTGGCTCAACAAGTCC	1113
	346	CysGlyGlyIleLeuProValArgAspArgProGlnIleLeuLeuGlySerAsnGlySer	365
QY	1114	CTTGGCCCTCATCTGATACAAACATCTTGCACATCCCGGCTCAAGTATTTGACCTGTG	1173
	366	ThrIleLeuThrLeuIleHisThrLeuLeuHisIleProThrGlnHisLeuTyrIleVal	385
QY	1174	GCTATAGTAATCTTCCATGACAAGAGCTCCCTGCAATTCGAGACAGACCTTGTGAG	1233
	386	ThrIysGlyTyrPheHisGlyLysHisSerProThrArgIleAspArgThrLeuLeuAsp	405
QY	1234	TTGCTTGAGATGTGTCTTTGTGGTCCCTGCACTGATCAGAGCTCGATATCAGAGAT	1293
	406	LeuPheGlyAspValPhePheValValProGlyLeuValThrAlaArgTyrHisAspArg	425
QY	1294	GCTGTGTGACCTGTCTATTTCTATGAGTTTGGGACCGGCTCAAGTGGTTTGAAGACAG	1353
	426	SerGlyGlyProValTyrPheTyrGlyIlePheGlnHisAspProHisCysPheGlnAsnSer	445
QY	1354	AAGCGGCTTTTGTCAAGCGGACGACGCTGATGAAGTCCGCTTTGTGTGGTGGTGGC	1413
	446	ArgProAlaPheValLysAlaAspHisThrAspArgGlyIleAspPheValPheGlyGlyPro	465
QY	1414	TTCTCGAAGGGGACATTGTTATGTTTCAGAGGACGACGAGAGAGAAAGTTACTGAC	1473

```

Db      466 PheLeuLySGLyAspValValMetHeGluGluAlaThGluGluGluLySLeuLeuSer 485
Qy      1474 CGGAAGATGATGAATAATCTAGGCGCTACCTTTGCTCGAACCGGGAATCTTAATGGAAACGAC 1533
Db      466 ArgLySmetMetLeuTyTrTrrpAlaAspPheAlaAArgSerGlyAspProbaGngLyAlaAsp 5050
Qy      1534 CTGTCTCTGTGCGCCAGCTTAAATCTGACTGAGCAGTACTCTCCAGCTCGAGCTTGAACATG 159
Db      506 LeuPProPrrTrpProValTyrAspGluTyrAspGluGlnTyrLeuGluLeuAspValaenille 5255
Qy      1594 AGCCCTCGGACAGAGACTCAAGAACCGCGGGTGGATTTTGTGGACCGACGACCATCCCTCTG 1655
Db      546 SerThrGlyAAspAspLeuAspValaGluPheTrThrAspThrLeuProLeu 5455
Qy      1654 ATCTGTGTGCTCTCCGACAGATGCTCCAGATGCTCTTCTTCTTAACTTTCTCTCTC 1711
Db      546 IleLeuSerAlaSerLyAlaLeuLeuSerProThrRheSerLeuIleLeuSerLeu 5655
Qy      1714 CTCACGCTTCTCTTTTCTTTTGTGCT 1740
Db      566 LeuSerProValLeuLeuSerAlaAla 574

RESULT 6
Q8I034 PRELIMINARY; PRT; 545 AA.
ID      08I034
AC      08I034;
DT      01-MAR-2003 (TREMBLrel. 23, Created)
DT      01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Carboxylesterase-like urinary excreted protein.
GN      Name=Cauxin;
OS      Felis silvestris catus (Cat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Flesipedida; Felidae; Felis.
OX      NCBI_TaxId=9685;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RA      MEDLINE=22458314; PubMed=12401131; DOI=10.1042/BJ20021446;
RX      Miyazaki M., Kamie K., Soeta S., Taira H., Yamashita T.;
RT      "Molecular cloning and characterization of a novel carboxylesterase-
RT      like protein that is physiologically present at high concentrations in
RT      the urine of domestic cats (Felis catus).";
RL      Biochem. J. 370:101-110(2003).
CC      -1- SIMILARITY: Belongs to the type-B carboxylesterase/1lipase family.
DR      EMBL: AB045377; BAC22577.1; -.
DR      HSSP; P12337; 1KAY.
DR      GO; GO:0016787; F:hydrolase activity; IEA.
DR      InterPro; IPR002018; CarboxylesteraseB.
DR      InterPro; IPR00379; Ser esters.
DR      Pfam; PF00135; Coesterase; 1.
DR      PROSITE; PS00122; CARBOXYLESTERASE B 1; 1.
DR      PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR      Hydrolase.
SQ      SEQUENCE 545 AA; 60505 MW; 9F73FA693D271FA9 CRC64;

Alignment Scores:
Pred. No.:      4,44e-139      Length:      545
Score:          1985.00      Matches:      373
Percent Similarity:      83.27%      Conservative:      60
Best local Similarity:      71.73%      Mismatches:      85
Query Match:      62.62%      Indels:      2
DB:          2      Gaps:      2

US-10-023-515-3 (1-1746) x Q8I034 (1-545)
Qy      91 GGGCTCTGTGTGTAAGGGCCACAGAGAAACACACAGGCTGGATGATTCAGGGCAAGCAA 1500
Db      25 GlyProAlaIaAlaAspAlaProValaArgSerThrArgLeuGlyTrrpAlaArgGlyLySGln 44
Qy      151 GTCACTGTGCTGGGAAGCCCTGTGCTGTGGAACGTGTTCTCTGGAGTCCCCCTTGTGCT 210

```



```

Db      45 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTyraIala 64
Qy      211 CCCCCTGGGAGATCCCTGGATTTAGAAACCCGACACCTGCATCCGCTCGGATAACTGG 270
Db      65 ProPheLeuGlyProLeuArgPheLeuGlnProLysProIalaLeuProLysAsnPhe 84
Qy      271 CGAAGACCACTCTTACCTTAATTTGTGCTCCAGAACTCAGAGAGTGGCTGTAGAT 330
Db      85 ArgAsnAlaThrSerTyProLysLeuCybPheGlnAspLeuGlnIlePheValSerTyx 104
Qy      331 CAACACATGCTCAGAGTGATTAACCGAAATTCGAGAGTCCAGAAAGCTGCTTACTG 390
Db      105 GlnHsValLeuLeuValArgTyProLysLeuGlnIalaSerGlnAspCybLeuTyx 124
Qy      391 AACATCTATGCGCTGCGCCAGCCGATTAAGAGCTCCAACTCCCGCTCTGTGGTGGTTC 450
Db      125 AsnIleTyraIaProIalaHsIalaAsnAsnGlySerAsnLeuProValMetValIlePhe 144
Qy      451 CCAGAGAGTGCCTTCAAGACTGGCTCAAGCTTCATCTTGTATGGTCCGCTCGGCTGCC 510
Db      145 ProGlyGlyAlaPheLeuMetGlySerAlaSerSerPheAspGlySerAlaLeuAlaIa 164
Qy      511 TATAGAGACGTCTGCTGTGTGCTCCAGTACCGGCTTGAAGAAATTTGGTTTCTTACC 570
Db      165 TyrgIuaAspValLeuIleValIleThrGlnTyraIeGlyIlePheGlyPhePheAsp 184
Qy      571 ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGGACAGAGTGGCTGCTGCC 630
Db      185 ThrGlyAspGlnHsIalaArgIleAsnIlePheAlaLeuAspGlnIalaIalaLeuThr 204
Qy      631 TGGGTCCAAGAAAGACATGAGTCTTCCGATGGGAAACCCAGCTGTGACCATTTTGGC 690
Db      205 TrpValaArgAspAsnIleGluPhePheGlyGlyAspProArgSerValIlePheGly 224
Qy      691 GAGTCCGGGGAGCCATAAGTCTTCTAGCTTATACGTCTCCATGGCCAAAGGCTTA 750
Db      225 GlnSerAlaGlyAlaIleSerAlaSerSerLeuIleLeuSerProIalaAsnIleLeu 244
Qy      751 TTCCCAAAAGCCATCATGAGAGTGGGAGTGGCCATCATCCCTTACCTG---GAGGCCAT 807
Db      245 PheHsIleValaIleMetGlnSerGlyAlaIleLeuProLeuLeuMetArgProPro 264
Qy      808 GATTATGAAAGAGTGAAGACTGCGAGTGGTGGCACTTCTGTGTGAACATGCGTCA 867
Db      265 GlyAspGlyArgIleValAspLeuGlnValIleAlaArgIleCybGlyCybHsIalaSer 284
Qy      868 GACCTGTAGGCGCTGCTGAGGCTGAGCAAAACCTCCAGAGACTGCTGACCTC 927
Db      285 AspSerAlaAlaLeuLeuGlnCybLeuArgAlaLysProSerGlnIleLeuMetAspIle 304
Qy      928 AGCCAGAAACAAAGTCTTTCATCTGAGTGGTGTGATGGTCTTCTTCTTAATGAGCCT 987
Db      305 SerIlyAspLeuThrPheSerIleProValIleAspAspPhePheProAspGluPro 324
Qy      988 CTAGATCTATTTGCTCAGAAAGCATTTAAAGCAATTTCTTCATCATCGAGTCAATAC 1047
Db      325 ValAlaLeuLeuThrGlnIleValaIlePheAsnSerValIleProSerIleIleGlyValaAsn 344
Qy      1048 CACAGAGTGGCTGCTGCTGCTGCTGAGAGAGGCTCTGAGTCTCAGAGGCTCCAAC 1107
Db      345 HisGlnCybAlaPheLeuLeu---SerThrIlePheSerGlnIleLeuGlyIleSerAsn 363
Qy      1108 AAGTCCCTTGGCTTCATCTGATCAAAACATCTGACATCCGCTCAGATATTGGAC 1167
Db      364 ArgSerIleuAlaLeuTyxIleValaHsIleThrPheLeuAsnIleProIleGlnIleuHs 383
Qy      1168 CTGTGGCTAATGAATACTCTTCATGACAGACACTCCCTGACTGAATCCGAGACGTCTT 1227
Db      384 LeuValaIalaAsnHsIleTyxPheTyxAsnIleHsSerProValGlnIleArgAspSerPhe 403
Qy      1228 CTGAGACTTGTGAGAGATGAGTCTTGTGGTCCCTGACATCCGCTCAGACATATCAC 1287
Db      404 LeuAspLeuLeuGlyAspValLeuPheValaValaIleProGlyValaIleThrAlaArgTyxHs 423

```

```

Qy      1288 AGAATGCTGTGTCACCTGTCTACTTCTATGAGTTTCGGACACCGGCTCACTGCTTGA 1347
Db      424 ArgAspAlaGlyAlaProValTyxPheTyxIlePheGlnHsProProGlnCybLeuAsn 443
Qy      1348 GACACGAAGCCGGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTGTGTGGT 1407
Db      444 AspThrArgProAlaPheValaIleAlaAspHsIleSerAspGlnIleArgPheValaPheGly 463
Qy      1408 GGTGCTTCTGAAAGGGGACATTTGATGTTCGAAGAGCCACGAGAGAGAAAGTTA 1467
Db      464 GlyAlaPheLeuTyxGlyAspIleValMetPheGlnGlyAlaIleThrGlnGlnIleuLeu 483
Qy      1468 CTGAGCCGGAGAGATGATGAATTAATCTGAGTCACTTGTGCAACCGGAATCCAAATGG 1527
Db      484 LeuSerArgIlyMetMetAspArgTyxIlePheAlaAsnPheAlaArgTyxIleAspProAsnIly 503
Qy      1528 AACGACCTGTCTGTGTGGCCAGCTTATTAATCTGACTGACAGTACCTGCACTGGA 1587
Db      504 GlnGlyValaProLeuIleProValaIleTyxIleThrGlnSerGlnGlnIleTyxLeuAspLeu 523
Qy      1588 AACATGAGCTTGGACAGAGACTCAAAAGACCGCGGCTGATTTTGGACCAACAATC 1647
Db      524 SerValSerValGlyGlnIleuLeuGlnGlnIleuValaGlnPheTyxMetAsnThrIle 543

RESULT 7
ID      095KH3      PRELIMINARY;      PRT;      361 AA.
AC      095KH3;
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Hypothetical protein.
OS      Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC      Cercopitheidae; Macaca.
OX      NCBI_TaxID=9541;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Temporal lobe right;
RA      Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA      Suzuki Y., Sugano S., Hashimoto K.;
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC      -I- SWIMLARTY: Belongs to the type-B carboxylesterase/lipase family.
CC      EMBL; AB060873; BAB46884.1; -.
DR      HSP; P12337; IK4Y.
DR      InterPro; IPR002018; CarboxylesteraseB.
DR      Pfam; PF00135; Coesterase; 1.
KW      Hypothetical protein.
SQ      SEQUENCE 361 AA; 41032 MW; AC62AB6E78C5F1A1 CRC64;

Alignment Scores:
Pred. No.:      3,74e-116      Length:      361
Score:      1676.50      Matches:      323
Percent Similarity:      96.47%      Conservative:      5
Best Local Similarity:      95.00%      Mismatches:      11
Query Match:      52.89%      Indels:      1
DB:      2      Gaps:      1

US-10-023-515-3 (1-1746) x 095KH3 (1-361)
Qy      724 ATACTGTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGATGGCC 783
Db      17 IleLeuSerProMetAlaGlnGlyLeuPheHsIleValaIleMetGlnSerGlyValaIa 36
Qy      784 ATCATCCCTTACCTGAGGCCATGATTAATGAGAAGAGTGAAGACCTGCAAGTGTGCA 843
Db      37 IleIleProTyxIleuTyxAlaTyxAspTyxGlnIlySerGlnAspLeuGlnValaIaIa 56
Qy      844 CATTTCGTGTGTAACAATGCGTCAAGACTCTGAGCGCTGTGAGGTGCTGAGGACAA 903
Db      57 HisPheCybGlyAsnAsnAlaSerAspSerGlnAspLeuAspArgCybLeuArgThrIlyx 76

```



```

OY 904 CCTTCAAGAGCTGTGACCTGAGCCGAGAAAACAAAGTCTTTCAGTCTGAGTGTAT 963
DB 77 SeisrlysgluuenuThleuSerGlnlyAlaIySerPheThrAlaGlyAlaSer 96
OY 964 GGTGCTTCTTCTTAATGAGCTCTAGATCTATTGTCTCAGAAAGATTAAAGCAAT 1023
DB 97 GlyAlaPhePheProAnGlnPheLeuGlnlyLeuSerGlnlyAlaPheGlyAlaIle 116
OY 1024 CTTTCATCATTCGAGGTGATTAACACGAGTGTGGCTTCTGCTGCTTATGAAAGAGCT 1083
DB 117 ProSerIleIleGlyAlaAlaAlaSerIleGlyGlyPheLeuLeuPheMet---GlnAla 135
OY 1084 CCTGAGATCTCAGTGGCTCCGAAAGTCCCTTGGCTCCATCTGATTAACAAACATCTG 1143
DB 136 ProGlnIleLeuSerGlySerAlaMySerIleAlaIleuIleGlnIleuIleLeu 155
OY 1144 CACATCCCGCTCAGATATTATGACCTTGTGCTAAATGATATCTTCATGACAGCACTCC 1203
DB 156 HistIleProGlnIleuIleuIleuIleuAlaAlaGlnIleuIleuIleuIleuIleu 175
OY 1204 CTGACTGAAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTCTTGTGGTCCCT 1263
DB 176 LeuThrGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 195
OY 1264 GCACGTGATCAGAGCTGATATCAGAGATGCTGAGACCTGCTACTCTATGATGTT 1323
DB 196 AlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 215
OY 1324 CGGACCGCGCTCAGTGTGCTTGAAGACAGCAAGCGGCTTGTGTAAGACCGCAAGCT 1383
DB 216 ArgHisArgProGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 235
OY 1384 GATGAAGTCCGCTTGTGTGCTGCTGCTGCTTCTGAAAGGGGACATTTATGTTGAA 1443
DB 236 AspGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 255
OY 1444 GGAGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1503
DB 256 GlyAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 275
OY 1504 GCTCGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1563
DB 276 AlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 295
OY 1564 GAGCAGTACCTTCAGCTGAGTGAATGAGCTTGAAGCTTGAAGCTGAAGCTGAAG 1623
DB 296 GlnGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 315
OY 1624 GTGATTTTGGACGAGACGATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
DB 316 ValGlnPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 335
OY 1684 CCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1743
DB 336 ProLeuSerPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 355

RESULT 8
SASB_ANAPL STANDARD: PRT: 557 AA.
AC 004791;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DB Fatty acyl-CoA hydrolase precursor, medium chain (EC 3.1.2.-)
DB (Thioesterase B).
OS Anas platyrhynchos (Domestic duck).
OC Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Amseriformes; Anacidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-65.
RC TISSUE=urophygial gland;

```

```

RX MEDLINE=93300823; PubMed=8314791;
RA Hwang C.-S., Kolatukudy P.E.;
RT "Molecular cloning and sequencing of thioesterase B cDNA and
RT stimulation of expression of the thioesterase B gene associated with
RT RT hormonal induction of peroxisome proliferation."
RL J. Biol. Chem. 268:14278-14284(1993).
CC -1- FUNCTION: Fatty acid biosynthesis chain termination and release of
CC the free fatty acid product is achieved by hydrolysis of the thio
CC ester by a thioesterase. This thioesterase may be associated with
CC peroxisome proliferation and may play a role in the production of
CC 3-hydroxy fatty acid diester phenomones.
CC -1- TISSUE SPECIFICITY: Highest levels in uropygial gland, much lower
CC in liver and kidney.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L05493; AAA9223.1; -.
DR PIR: A47162; A47162.
DR HSSP: O7540; 1K4Y.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR00379; Ser_estr.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE B_2; 1.
KW Direct protein sequencing; Fatty acid biosynthesis; Hydrolyase; Signal.
FT SIGNAL 1 25
FT CHAIN 26 557 Fatty acyl-CoA hydrolase.
FT ACT_SITE 227 227 Acyl-ester intermediate (By similarity).
FT ACT_SITE 345 345 Charge relay system (By similarity).
FT ACT_SITE 460 460 Charge relay system (By similarity).
FT DISULFID 93 122 By similarity.
FT CARBOHYD 476 476 N-linked (GlcNAc...) (potential).
SQ SEQUENCE 557 AA; 61637 MW; 03B35D90A037F680 CRC64;

Alignment Scores:
Pred. No.: 6,456-88 Length: 557
Score: 1298.50 Matches: 265
Percent Similarity: 62.77% Conservative: 79
Best Local Similarity: 48.36% Mismatches: 177
Query Match: 40.96% Indels: 27
DB: 1 Gaps: 8

US-10-023-515-3 (1-1746) x SASB_ANAPL (1-557)
OY 88 ACTGGGCTTCTGCTGAAGGCGCAGAGAACACAGGCTGGATGATTCAGGCGAAG 147
DB 23 ThrGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 42
OY 148 CAAGTCACGTGCTGGAAGCCCTGCTGGAACGCTGCTGGAAGCTGCTGCTGCTGCT 207
DB 43 GlnValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 62
OY 208 GCTCCCGCTGGATTCCTGATTTACGAACCCGAGCCCTGATCCCTGGAGTAAC 267
DB 63 LysProProValGlyProLeuIleuIleuIleuIleuIleuIleuIleuIleuIleu 82
OY 268 TTGCGAGAGCCACTCTTACCTTAATTTGCTGCTCAGAACTCAGAGTGGCTCTTGA 327
DB 83 ValArgAspAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 99
OY 328 GATCAACAGATGCTC-----AAGTGCAATTACCCGAATTCGAGATGCA 372
DB 100 GlnGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 119
OY 373 GAAGACTGCTCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432

```

Db 120 GluAerCyLeuLeuValTyrThrProValSerThrGluGluGluValSer 139
 QY 433 CCGGTGTTGGTGGTCCCAAGAGGTGCCTTAACAGCTGCACCTCCATCTTGAT 492
 Db 140 ProValPheValITripLeuHISglYglYglYLeuValSerGlyValAlaSerSerYrAap 159
 QY 493 GGGTCGCCCTGGCTGCTCCATGAGAGAGAGCTGGTGTGGTGGTCCAGTACCGGCTAGGA 552
 Db 160 GlySerAlaLeuAlaIaIaPheAapAenValValValThrIleGlnYrHrGluGly 179
 QY 553 AATTTGGCTTTCTTCAACACATGGAGTACAGCATGCTCCGGGGAATGGGCTTCAAGAC 612
 Db 180 IleaIaGlyYrPheSerThrGlyAapYrSHIaIaArgGlyAaSnTrpYrLeuAap 199
 QY 613 CAGGTGGCTGCTGTGCTCCGAGAGACATCGAATGTTCTGGTGGGAGCCCAAGC 672
 Db 200 GlnValAlaAlaLeuGlnITripIleGlnGluAaSnITrIleHISpHeArgGlyAapProGly 219
 QY 673 TCTGTGACCATCTTTGGCAGATCCGGGGAGCCATAGTGTCTTACTTACTGTCT 732
 Db 220 SerValIThrIlePheGlyGluSerAlaGlyValaIseValSerAlaLeuValLeuSer 239
 QY 733 CCGATGGCCCAAGGCTTATTCACAAAGCATATGAGAGTGGGGTCCATC---ATC 789
 Db 240 ProLeuAlaYrGlyLeuPheHISYrAlaIleSerGluSerGlyThrAlaValArgIle 259
 QY 790 CCTTACTGAGAGCCCATGATTATGAGAAAGTAGAGAGACCTCGAGTGGTGCACATTTTC 849
 Db 260 LeuPheThrGlu-----GlnProGluGluGlnAlaGlnArgIleAlaAla 275
 QY 850 TGTGTTAAACAATGCTCAGACTCTGAGAGCTGCTGAGGTGCTGAGAGCAAAACCTCC 909
 Db 276 AlaGlyCySgluYrSerSerSerAlaAlaLeuValGluCyLeuArgGluYrThrGlu 295
 QY 910 AAGGAGCTG-----CTGACCTTCAGCCAGCAAAACAAAGCTTTTACTCGAGTGTGAT 963
 Db 296 AlaGluMetGluGlnIleThrLeuYrSwetProPomePheIleSerAlaSerLeuAap 315
 QY 964 GGTGCTTCTTCTTCTAATGAGGCTCTAGATCTATGTCTCAGAAAGCATTTAAACAAAT 1023
 Db 316 GlyValPhePheProYrSerProArgGlnLeuLeuSerGluYrValIleAaSnAlaVal 335
 QY 1024 CTTTCCATCATCGAGTCAATTAACACAGAGTGTGCTTCTGCTGCTCCT---ATGAAGAG 1080
 Db 336 ProYrIleIleGlyValaIaAaSnCySgluPheGlyTrpIleLeuProArgMetCetYrS 355
 QY 1081 GCTCCTGAGATCCCTCAGTGCCTCCAAACAAGTCCCTTGCCTCCATCTATCAAAACATC 1140
 Db 356 PheProGluPheThrGluGlyLeuGluYrAapValaIaArgGlnValLeuGlnSerThr 375
 QY 1141 CTGCACATC-----CCGCTCAGTATTGCACTTGTGGCTAATGAATAC 1185
 Db 376 LeuAlaLeuSerPheYrGlyAlaIeProSerAapIleValaIeAapLeuValTyrAaSnITrYr 395
 QY 1186 TTCCATGACAAACATCCCTGATGAATCCGAGACAGTCTTGTGAGCTTGTGAGAT 1245
 Db 396 IleGlyValaIaGluAaSnArgAlaGlnValaIaArgAapGlyLeuLeuAapSerIleAlaAap 415
 QY 1246 GTGTCTTGTGGTCCCTGAGCATATGATCAGCTGATGATCAAGAGCTGGTGCACCT 1305
 Db 416 ProLeuPheValIPheSerAlaValaIaGluValaIaArgHISAlaArgAapAlaGlyAaSnTr 435
 QY 1306 GTCTACTTCTATGAGTTCGAGACCGGCTCAGTGTGGAAGACACAGACCGGCTTTT 1365
 Db 436 ValYrPheYrGluPheGlnHISArgProSerSerAlaAlaGlyValaIaValProGluApe 455
 QY 1366 GTCAAAGCCGACCGCTGATGAAGTCCGCTTGTGGTGGTCCCTTCTGAAGGGG 1425
 Db 456 ValYrValaIaAapHISAlaAapGluIleAlaPheValIPheGlyYrProPheLeuAlaGly 475
 QY 1426 GACATTTGATTGTTGCAAGAGCGACGAGGAGGAAGTTCAGTACCGGCGGAATGATG 1485
 Db 476 AaSn-----AlaThrGluGluGlnAlaYrLeuSerArgThrValMet 489

QY 1486 AAATACGGCTACCTTTGCTCGAACCCGGGAATCTTAATGGGAACGACCTGTCTGTGG 1545
 Db 490 LysYrITrPThrAaSnPheAlaArgAaSnIaAaSnProAaSnGlyGluGlyLeuValaIHISTrp 509
 QY 1546 CCAGCTTAAATCTGACTGAGACGAGTACCTCCAGCTGAGTGAACATGAGCTCGAGACG 1605
 Db 510 ProGlnYrAapMetAaGluArgYrLeuGlnIleAaPleuThrGlnIleValaIaYrS 529
 QY 1606 AGACTCAAAAGAACCGCGGTGAGATTTTGGACCAAGACCATCCCTGATCTGTGTGCC 1665
 Db 530 LysLeuYrSgluAArgYrMetGluPheTrpMetGlnLeuThrGluGlnIleMetSerAap 549
 QY 1666 TCCGACATGCTCCACAGTCTCTT 1689
 Db 550 ArgArgArgYrGlyHISThrAapLeu 557
 RESULT 9
 Q91WG0 PRELIMINARY; PRT; 561 AA.
 AC Q91WG0:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Carboxylesterase 2 (Carboxylesterase M13 precursor).
 GN Name=Ces2; Synonyms=Ces2A3;
 OS Mus musculus (Mouse);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TissuS=Colon; and Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carrinci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fealey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smalley D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TissuS=Kidney;
 RA Strauberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TissuS=Colon;
 RA Strauberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TissuS=Colon;
 RA Strauberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TissuS=Colon; and Kidney;
 RA Strauberg R.;


```

OY 211 CCCCCGCGGGATCCCTGGATTTTACGAACCCGAGCCCTGATGCCCTGGGATAACTTG 270
DB 66 ProProValIGlyProLeuMetArgPheAlaProProGluAlaProGluProThrSerGlyVal 85
OY 271 CGAAGAGCCACCTCTTACCTTAATTTGTGCTCCAGAACTCAGATGGCTG--CTCTTA 327
DB 86 ArgArgGlyThrSerHisProLeuMetCysLeuGlnAlaAspAlaSerLeuMetGlySerGlu 105
OY 328 GATCAACACATGCTCAAGGTCATTACCCGAAATTCGAGTCTCGAAGATGCTCTTAC 387
DB 106 AspLeuLysMetMetAspLeuLysLeuProProLysSerMetSerGlyAspCysLeuThr 125
OY 388 CTGAACATCTATGCGCTGCCAGCCGATACAGGCTCAAGCTCCCGCTGGTGGTGG 447
DB 126 LeuAsnLysLeuValProAlaHisAlaHisGlyGlySerAspLeuProAlaMetValThr 145
OY 448 TTCCGAGAGGTCCTTCAAGACTGCTCAGCTTCATCTTGTATGGTCCGCGCTGGCT 507
DB 146 IleHisGlyGlyAlaLeuThrValGlyMetAlaSerMetGlyAspGlySerMetLeuAla 165
OY 508 GCTATGAGAGACGTGCTGCTGCTGCTGCTCAGTACCGGCTTGAAGAAATTTGTTTCTTC 567
DB 166 AlaThrGluAspValValValValAlaIleGlnThrArgGlyGlyValLeuGlyPhePhe 185
OY 568 ACCCATGAGGATCAGACATGCTCCGGGGAATGGGCTTCAAGACACAGGTGGCTCTTG 627
DB 186 SerThrGlyAspGlnHisAlaLysGlyAsnThrGlyThrLeuAspGlnAlaAlaLeu 205
OY 628 TCTTGGTTCAGAAAGATCAGATGATTTCTCGTGGGAGCCGACGCTGTGTGACATCTTT 687
DB 206 ArgThrValGlnGlnAsnLysIleValHisPheGlyGlyAsnProAspArgValThrIlePhe 225
OY 688 GCGGAGTCCGCGGAGCCATAAGTTTCTAGTCTTATATCTGCTCCATGGCCAAAGCC 747
DB 226 GlyGlnSerAlaGlyGlyThrSerValSerSerHisValValSerProMetSerGlnGly 245
OY 748 TTATTTCCCAAGCATATGATGAGTGGGTCGATCATCCCTTACCTGAGGCCAT 807
DB 246 LeuPheHisGlyAlaIleMetGlnSerGlyValAlaValLeuProAspLeuLysSer 265
OY 808 GATTATGAGAGAGTGAAGACTGAGGATGTTGACATTTTGTGTGAACATCGTCA 867
DB 266 SerSerGluMetValHis-----ArgIleValAlaAsnLeuSerGlyCysAlaAlaVal 283
OY 868 GAGTCTAGGCGCTCTGAGTGGCTGAGGACAAACCTCCAGAGACTGCTGACCTC 927
DB 284 AsnSerGluThrLeuMetCysCysLeuArgGlyLysAsnGluAlaGluMetLeuAlaIle 303
OY 928 AGCCGAGAAACAAAGTCTTCACTGAGTGTGATGTGCTTCTTCTTAATAGCTT 987
DB 304 AsnLysValPheLysIleIleProGlyValAlaAspGlyGluPheLeuProLysHisPro 323
OY 988 CTAGATCTATTGTCAGAAAGCATTTAAAGCAATTCCTTCATGATCGAGTCAATAAC 1047
DB 324 GlnGlnLeuMetAlaSerLysAspPheHisProValProSerIleIleGlyIleAsnAsn 343
OY 1048 CACGAGTGGCTTCTGCTGCT----- 1071
DB 344 AspGluThrGlyThrIleLeuProThrIleLeuAspProAlaGlnLysIleGluGluIle 363
OY 1072 ---ATGAAGAGAGGCTCTGAGATCTCAGTGGCTCCAAACATCCCTTGCCTCATCTG 1128
DB 364 ThrArgLysThrLeuProAlaValLeu-----LysSerThrAlaLeuLysMet 379
OY 1129 ATACAAACATCTGACATCCCGGCTCAGATTATTTGACCTTGAGGCTAATGATCTTC 1188
DB 380 Met-----LeuProProGluCysGlyAspLeuLeuMetGluGluThrMet 394
OY 1189 CATGCAACGACTCCCTGATCTGAATTCGAGACAGTCTTCTGAGACTGCTGAGAGATG 1248
DB 395 GlyAspThrGluAspProGluThrLeuGlnAlaGlnPheArgGluMetCysGlyAspPhe 414

```

```

OY 1249 TTCTTTGTGTCCTCCATCATATGATCAGCTCATATGACAGATGCTGTGCACTGTC 1308
DB 415 MetPheValIleProAlaLeuGlnValAlaHisPheGlnArg---SerHisAlaProVal 433
OY 1309 TACTTCTATGAGTTTGGGACCGGCTCTCAGTGTCTTGAAGACAGACCGGCTTTTGTTC 1368
DB 424 TyrPheThrGluPheGlnHisArgProSerPhePheLysAspPheArgProProThrVal 453
OY 1369 AAAGCCGACACCGCTGATGAATCCGCTTGTGTGCTGGTGGCTCTTCTTGAAGGGGAC 1428
DB 454 LysAlaAspPheLysGlyAspGluIlePheLeuValPheGlyThrGlnPhe-----GlyAsn 471
OY 1429 ATTGTATGTTTCAAGAGACGACGAGGAGAGAGATTAATGAGCCGAGATGATGA 1488
DB 472 IleLysLeu-----ProThrThrGluGluGlnGlnLeuSerArgArgIleMetLys 489
OY 1489 TACTGGCTTACCTTGTCTCGAACCGGAAATCTTAATGGAACACCTGTCTGTGGCCA 1548
DB 490 TyrThrAlaAsnPheAlaArgHisGlyAsnProAsnSerGluGlyLeuProThrThrPro 509
OY 1549 GCTTATATGCTGACCTGACGATCCTCCAGCTGACCTTGAACATGAGCCTGAGACAGA 1608
DB 510 ValMetAspPheLysAspGluGlnThrLeuGlnLeuAspIleGlnProSerValGlyArgAla 529
OY 1609 CTCAAGAACCGGAGTGAATTTTGGACACGACCATCC 1650
DB 530 LeuLysAlaArgGluGlnPheThrThrThrLeuPro 543

RESULT 11
O35533 PRELIMINARY; PRT; 559 AA.
ID O35533
AC O35533
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Carboxylesterase precursor (EC 3.1.1.1).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian golden; TISSUE=Liver;
RA Sone T., Wang C.Y.;
RT "Microsomal amidases and carboxylesterases.";
RL (in) Gengerich F.P. (eds.),
RL COMPREHENSIVE TOXICOLOGY VOLUME 3, Biotransformation, pp.265-281,
RL Pergamon, Oxford (1997).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian golden; TISSUE=Liver;
RA Sone T., Ishida Y., Takabatake E., Wang C., Isobe M.;
RL Submitted (May-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL, D50577; BA023605.1; -.
DR HSBP, P12377; IKAY.
DR GO, GO:0004091; F:carboxylesterase activity; IEA.
DR GO, GO:0016787; F:hydrolase activity; IEA.
DR InterPro, IPR002018; Carboxylesterase.
DR InterPro, IPR000886; ER_target.S.
DR InterPro, IPR000379; Set_estra.
DR PROSITE, PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE, PS00941; CARBOXYLESTERASE_B_2; 1.
DR PROSITE, PS00014; ER_TARGET; UNKNOWN_1.
KW Hydrolase; Signal.
FT SIGNAL 1
FT CHAIN 28 559 potential.
SQ SEQUENCE 559 AA; 62126 MW; 5B68D4F9A47C6AD8 CRC64;
carboxylesterase.

Alignment Scores: 2.82e-86 Length: 559
Pred. No.: 1276.50 Matches: 262
Score:

```

Percent Similarity: 64.58% Conservative: 79
 Best Local Similarity: 49.62% Mismatches: 172
 Query Match: 40.27% Indels: 15
 DB: 2 Gaps: 9

US-10-023-515-3 (1-1746) x 035533 (1-559)

QY 82 GAAAAAAGCTGGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAG 141
 DB 26 GtyninabpserProser-----ProileghantthiethGlycylValArg 42
 QY 142 GCGAAGCAAGTCAGTCTGTGGAAAGCCCTGTGCTGAACGTGTCTCGAGTCC 201
 DB 43 GlyArgLeuValHisValLysAspThrGluValAspValHisLeuThrPheLeuclYllePro 62
 QY 202 TTGTGTGCTCCCGCTGGATCCCTGGATTAAGAACCCGACCTGCATCGCCCTGG 261
 DB 63 PheAlaLysProProValGlyProLeuArgPheAlaProProGluProProGluProTrp 82
 QY 262 GATAACTGGAGAAAGCCCTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTG 321
 DB 83 SerGlyValArgAspGlyAsnSerPheProAlaMetCysLeuGlnAspAspMetMet 102
 QY 322 CTCTTAGAT---CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGCAGAGAC 378
 DB 103 AsnSerGluGlyLeuLysMetCileuLysLeuileMetProProIleSerMetSerGluAsp 122
 QY 379 TGCCCTTACCTGAACATCTATGCGCCGCCACGCGCGATACAGCTCCAGTCCCTG 438
 DB 123 CysLeuLysLeuAsnLetyrThrProThrHisAlaHisGluGlySerAsnLeuProVal 142
 QY 439 TTGTGTGCTCCAGAGAGTCCCTTCAAGACTGCTCAGCTCCATCTTGTGATGGTCC 498
 DB 143 MetValTrpIleHisGlyGlyAlaLeuValValGlyMetCilaSerMetLysAspGlySer 162
 QY 499 GCCCTGCTGCTATGAGAGCTGCTGTGTGTGCTGCCTCAGTACCCGCTTGAATATT 558
 DB 163 MetLeuAlaAlaIleGluAspValValValThrIleGlnLysArgLeuGlyAlaLeu 182
 QY 559 GGTTCCTTACACACATGGGATCAGATGCTCCGGGAACTGGGCTTCAAGACCAAGGTG 618
 DB 183 GlyPhePheSerThrGlyAspGluHisAlaArgLysAsnTrpGlyLysLeuAspGlnAla 202
 QY 619 GCTGCTGTGCTGGGTCCAGAAAGACATGATTTCTGCTGGAGCCCAAGCTCTGTG 678
 DB 203 AlaAlaLeuArgTrpValGlnGlnAsnIleAlaHisPheGlyGlyAsnProGlyGlnAla 222
 QY 679 ACCATCTTGGCGAGTCCGCGAGCCATAAGTGTCTTACTTACTGTCTCCATG 738
 DB 223 ThrIlePheGlyGlnSerAlaGlyGlyLysSerValSerSerHisValValSerProMet 242
 QY 739 GCCAAAGGCTTATTCACAAAGCATGATGAGAGTGGGGTGGCATGATCCCTTAACTG 798
 DB 243 SerLysGlyLeuPheHisGlyAlaIleLeuGlnSerGlyValSerValMetProGlyLys 262
 QY 799 GAGGCCCATGATTTGAGAAAGTGAAGAGCTGCAGTGTGTGACATTTTGTGTGTAAC 858
 DB 263 IleSerSerSerSerGluMetVal-----TyrThrIleValAlaAsnLeuSerAspGly 280
 QY 859 AATCGCTCAGACTCTGAGGCTGTGCTGAGTGTGCTGAGAGCAAAACCTCCAGAGCTG 918
 DB 281 AlaAlaValAsnTrpGluThrLeuValSerCysLeuArgGlyLysSerGluAlaGluIle 300
 QY 919 CTGACCCCTGAGCGAGAAACAAAGCTTTCAGTGCAGAGTGTGATGGCTTTCTTCTCT 978
 DB 301 LeuAlaIleAsnLysValPheLysIleIleProAlaValValAspGlyGluPheLeuPro 320
 QY 979 AATGAGCTCTAGATCTATTGTCTCAGAAACATTTAAAGCAATTCCTTCATCATCGGA 1038
 DB 321 LysHisProGlnGlnLysLeuValSerLysLysAspPheAspProValProSerIleIleGly 340
 QY 1039 GTCAATATACCAAGGTGTGCTTCTGCTGCT-----ATGAAAGAGCTCTCGAGATC 1092

DB 341 ValAsnAsnAspGluLysGlyTrpLeuValProValIleMetGlySerAlaGlnGluIle 360
 QY 1093 CTCAGTGGCTCCACAAAGTCCCTT-----GCCCTTCATCTGATACAAACATCTCGAC 1146
 DB 361 LysGluIleThrArgValThrLeuProAlaIleLeuLysSerThrAlaAsnGlnMetLys 380
 QY 1147 ATCCGCTCAGATTTTTCAGCTTGTGCTAATGAAATCTTCATCCATACAAACATCCCTG 1206
 DB 381 LeuProProGluLysGlyLysAspLeuMetGluGluLysTrpMetGlyAspAsnGluAsp 400
 QY 1207 ACTGAATCCGAGACAGTCTTCTGACTTGTGAGATGTGTTCTTGTGCTCCGCA 1266
 DB 401 GlnThrLeuAlaHisGlnPheArgGluMetMetGlyAspPheThrPheValIleProSer 420
 QY 1267 CTGATCAGAGTGCATATCACAGAGATGCTGCTGCTCTCTACTTATAGATTGG 1326
 DB 421 LeuGlnValAlaHisPheGlnArg---AlaHisAlaProValTyrPheLysGluPheGln 439
 QY 1327 CACCGGCTCAGTCTTGTGAAGACAGAGCGGCTTTGTCAAGCGGACACAGCTGAT 1386
 DB 440 HisArgProSerPhePheLysAspThrLysProProHisValLysAlaAspHisGlyAsp 459
 QY 1387 GAAGTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGACATGTTATGTTGAAAGCA 1446
 DB 460 GluPhePhePheValPheGlyAsn---LeuLeuLysGly-----IleLysPhe---Ala 475
 QY 1447 GCCACGAGAGAGAGAGTATCTGAGCCGAGATGAAATGATCTGGGCTGCTTGTCT 1506
 DB 476 SerThrGluGluGluGluLeuLeuSerArgLysMetCysLysTrpAlaAsnPheAla 495
 QY 1507 CGAACCGGAGATCTTAATGGGAAAGACCTGTCTGTGCGCCAGCTTAACTGATGAG 1566
 DB 496 ArgHisGlyAsnProAsnSerGluAsnLeuProLysTrpProMetLeuAspHisAspGlu 515
 QY 1567 CAGTACTCCAGCTGACATTGAACATGAGCTCGACAGAGACTCAAGAAACCCGGGTG 1626
 DB 516 GlnTrpLeuGlnLeuAspIleLysProAlaValAlaGlyArgAlaLeuLysValArgArgLeu 535
 QY 1627 GATTTTGGACACAGACATCCCTCC 1650
 DB 536 GlnPheThrLysLysIleLeuPro 543

RESULT 12
 Q8BM97 PRELIMINARY; PRT; 559 AA.
 AC Q8BM97;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE Mus musculus adult male colon cDNA, RIKEN full-length enriched
 DE library, clone:9030624J02 product:similar to CARBOXYESTERASE (EC
 DE 3.1.1.1) (A1-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
 DE (PROCAINE ESTERASE) (METHYLBUTYRASE).
 GN Name=9030624J02Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=92279253; Pubmed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
 RU [2]
 RP SEQUENCE FROM N.A.
 RC STAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=21085660; Pubmed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
 RU [3]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20493974; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Saeki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Kikunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasai H.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara B., Wachihi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA Adachi J., Atzawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Nagaoka T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sato H.,
RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AK033563; BAC28361.1; -
DR HSSP: P12337; 1K4Y.
DR MGI: MGI:2443170; 9030624L02R1k.
DR GO: GO:0005615; C:extracellular space; TAS.
DR InterPro: IPR002018; Carboxylesterase.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 559 AA; 62305 MW; 8C54D21522C252C4 CRC64;

Alignment Scores:
Pred. No.: 7,24e-86 Length: 559
Score: 1271.00 Matches: 260
Percent Similarity: 63.30% Conservative: 78
Best Local Similarity: 48.69% Mismatches: 166
Query Match: 40.09% Indels: 30
DB: 2 Gaps: 8

US-10-023-515-3 (1-1746) x Q8BM97 (1-559)
QY 91 GGGCTTCTGCTGAAGGCGGACAGACACAGGCTGGATGATTCAGGCGCAAGCA 150
DB 26 GAGTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 45
QY 151 GTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 210

DB 46 ValHisValIysAspThrAspIleAlaValHisThrPheLeuGlyIleProPheAlaIys 65
QY 211 CCCCCGCTGGATCCCTTCGATTTACGAACCCGACGCTGCATCCCTGGATTAAGT 270
DB 66 ProProValIleGlyProIleuAlaPheAlaProProGluAlaProGluProIleSerIleGly 85
QY 271 CGAAGAGCCACCTCTTACCTTAATTGGCTCCGAGACTCGAGATGAGTGGCTG--CTCTTA 327
DB 86 ArgAspGlyThrSerHisProAsnMetCysLeuGlnAsnAspAsnLeuMetCysIleGlu 105
QY 328 GATCAACACATGCTCCAGTCCGATTAACCGAATTCCGAGTCCGAGAACTGCTTAC 387
DB 106 AspLeuIysMetMetAsnLeuIleLeuProProIleSerMetSerGluAspSerLeuIyr 125
QY 388 CTGAACATCTATGCGCCCTGCCCGCCATACAGGCTCCAGCTCCCTCTGCTGTCG 447
DB 126 LeuAsnIleIyrValProAlaHisAlaHisIleGluIysSerAsnLeuProValMetValTyr 145
QY 448 TTCCGAGAGGCTTCGACAGCTGCTCAGCTCCATCTTGAATGGGTCGCGCTGGCT 507
DB 146 IleHisIleGlyAlaLeuThrValGlyMetAlaSerMetIyrAspGlySerMetLeuAla 165
QY 508 GCGTATGAGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
DB 166 AlaThrGluAspValValValAlaValAlaIleGlnIyrArgLeuGlyValLeuGlyPhePhe 185
QY 568 ACCACATGAGATCAGACATGCTCCGCGGAACTGGGCTCCAGACACAGCTGCTGCTG 627
DB 186 SerThrIyrAspGlnHisAlaIleAlaIleValSerIyrAsnIyrGlyIyrLeuAspGlyValAlaIleLeu 205
QY 628 TCCGTGCTCAGAGAAACATGAGTCTTGGTGGGAGCCCGACGCTCTGTGACATCTT 687
DB 206 ArgIyrValGlnGlnAsnIleValHisPheGlyGlyAsnProAspArgValThrIlePhe 225
QY 688 GCGGAGTCCGCGGAGCCATAGTCTTCTAGTCTTATGATGCTGCTCCATGCGCAAGGC 747
DB 226 GlyIleSerIleIleGlyIleThrSerValSerSerHisValIleSerProMetSerGlnGly 245
QY 748 TTATTCACAAAGCCATGAGAGAGTGGGCTGATCATGCTTACCTGAGAGCCCAT 807
DB 246 LeuPheHisIleGlyAlaIleMetGluSerIleValAlaValIleProAspLeuIleSerSer 265
QY 808 GATTATGAGAGAGTGAAGACCTCGAGGTGGTGGACATTTCTGTGTAACAATGCTGCA 867
DB 266 SerSerGluMetValHis-----ArgIleValAlaAsnLeuSerGlyCysAlaIleVal 283
QY 868 GACTGTGAGGCTGCTGAGGTGCTGAGCAAAACCTCCAGAGCTGCTGACCTG 927
DB 284 AsnSerGlyThrIleuMetCysCysLeuArgIleGlyAsnGluAlaGluMetLeuAlaIle 303
QY 928 AGCCAGAAACAAAGCTCTTCACTCGAGTGGTGAATGCTGCTTCTTCTTAATAGCTT 987
DB 304 AsnIysValPheIysIleIleProGlyValAlaAspGlyGluPheLeuProIyrHisIlePro 323
QY 988 CTAGATCTATGCTTCAGAAAGCATTAAGCAATTCCTTCATGATGAGTCAATAC 1047
DB 324 GlnGluLeuMetAlaSerIysAspPheHisProValProSerIleIleGlyIleAsnAsn 343
QY 1048 CACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071
DB 344 AspGlyIyrGlyIyrIlePheLeuProThrIleMetAspProAlaGlnIleIleGluIle 363
QY 1072 ---ATGAAGAGGCTCTGAGATCTCAGTGGCTCCAAAGTCCCTGCTTCATCTG 1128
DB 364 ThrArgIyrThrIleuProAlaValIleu-----LysSerThrAlaLeuIysMet 379
QY 1129 ATACAAACATCTGCAATCCCGCTCAGTATTTGCACTTGGTGGCTAAGATATCTTC 1188
DB 380 Met-----LeuProProGluIyrGlyAspLeuLeuMetGluIleIyrMet 394
QY 1189 CATGACAAGACCTCCCTGATCGAATCCGAGACAGTCTTCTGAGCTTCTGAGATGTG 1248


```
Dh 395 GlyaerthrgluaspProglutrhreuglnalaglInpheargluMetlysglyasphe 414
Qy 1249 TTCTTGGTCCCGCAGCTGATCAAGCTGATATCAAGAGAGTGGTGGTCACTGTC 1308
Db 415 MetpnevalileproalaleuqInValAlahlspehlnary--Serihlsalproval 433
Qy 1309 TACTTCTAGAGTTTCGGACCGGCTCGATGCTTTGGAAGACAGACGCGCTTTTTC 1368
Db 434 TyrPheTygluPheglInhlsarProserPhePheTyaspPheargProtyrval 453
Qy 1369 AAAGCCGACCGCTGAGTAATCCGCTTGTCTGGTGGTGGCTTCTGTAAGGGGAC 1428
Db 454 LysAlaAspPhlsGlyAspGluIlePheLeuValPheglTyrglnphe-----Glyasn 471
Qy 1429 ATGTGTTATGTTGGAAGGACCGAGGAGAGAGTAATCTGAAGCGGAGATGATGAA 1488
Db 472 Ilelyleu-----ProtyrThrIugluIugluInleuSerArgarghewelys 489
Qy 1489 TACTGGGCTACCTTGTCTGGAACCGGAATCTTAATGGGAACGACTGTCTGTGGCCA 1548
Db 490 TyrArgAlaAspPhlsArgHlsGlyAspProAsnserGlygluLeuProtyrtrpPro 509
Qy 1549 GCTTATATCTGACGACGACGATCTCCAGCTGGACTTGAAACATAGCTCGACAGAGA 1608
Db 510 ValMetAspPhlsAspGluIuglnTyrgleuGlnleuAspIleGlnProserValglyArgAla 529
Qy 1609 CTCAAGAACCGCGGTGATTTTGGACGACGACCATCC 1650
Db 530 LeuValAlaArgArgleuGlnPheTrpIlysrIleuPro 543

RESULT 13
Q6PDB7 PRELIMINARY; PRT; 556 AA.
ID Q6PDB7
AC Q6PDB7
DT 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein BC015286.
GN Name=BC015286;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Franke C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.I., Skalske U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strauberg R.;
RA Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/1lipase family.
```

```
DR EMBL: BC058815; AAH58815.1; -.
DR HSSP: P12337; 1k4Y.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR002018; CarboxesteraseB.
DR InterPro: IPR003579; Ser_eatrs.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
DR KX Hydrolyase; Hypothetical protein.
SQ SEQUENCE 556 AA; 61927 MW; 58DB05A3C73BEDC CRC64;

Alignment Scores:
Pred. No.: 1,56e-85 Length: 556
Score: 1266.50 Matches: 273
Percent Similarity: 60.17% Conservative: 76
Best Local Similarity: 47.07% Mismatches: 190
Query Match: 39.95% Indels: 41
DB: Gaps: 12

US-10-023-515-3 (1-1746) x Q6PDB7 (1-556)
Qy 1 ATGCCACAGGAGCTTACTTCAATCTGCTTCAAGATGGCTTTTCGATT-----CTC 54
Db 1 MetProArg-----SerGlnMetHlsAsnTrpIleuAspValIleuLeuPheGlyLeu 17
Qy 55 CAGCCCTGTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGGAAGGSCACAG 114
Db 18 LeuLeuLeuLeuGlyHlsValGln---GlyGlnAspSerProGluAla---SerProIle 35
Qy 115 AGGAACACACAGCTGGAGATGATTCAGAGGCACAGACATCACTGTCTGGGAAGCCTGTG 174
Db 36 ArgAsnThrHlsThrIleGlnValAlaArgGlySerLeuValHlsValIlyAspThrIlyAla 55
Qy 175 CTTGTGAACGTGTTCTCTCGAGATCCCTTTGCTGCTGCTCCCGCTGGAGATCCCTGGATT 234
Db 56 GlyValHlsThrPheLeuGlyIleProPheAlaIlyAspProValIglyProLeuAlaArgPhe 75
Qy 235 ACGAACCCGACCTGGATGCGCTGGGATTAATCACTGAGAGAGACACCTGCTTAAT 294
Db 76 AlaProProGluAlaIarProGluProItrPserGlyValAlaArgAspGlyThrAlaHlsProAla 95
Qy 295 TTGTGCTTCCAGAACTCAGAGTGGCTGCTTTAGATCAACACATGCTCAAGTGGCATTAC 354
Db 96 MetCysLeuGlnAsn-----LeuGlyValMetIyAspGlyIleIyLeuIyLeu 111
Qy 355 CCGAAATTCGAGAGTGCAGAAACATGCTGCTAAGCTTAACATTAATGCGCTGCGCCACAGCC 414
Db 112 ProProValSerThrSerGluAspCysLeuIyLeuAsnIleTyTrnProAlaHlsAla 131
Qy 415 GATACAGAGCTCAAGCTCCCGCTGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGGC 474
Db 132 HlsGluIySerHlsLeuProValMetValTrpIleHlsGlyGlyLeuValAlaGly 151
Qy 475 TCAGCTTCATCTTTGATGGGTCGCGCTGGCTGCTGATAGAGACGCTGGTGTGCTG 534
Db 152 MetAlaSerMetTyrgAspGlySerLeuLeuAlaIleGluAspLeuValIValThr 171
Qy 535 GTCCAGTACCGGCTAGGAATATTGTTGTTCTTCAACACATGAGTACGATGCTCGCGGG 594
Db 172 IleGlnTyrgleuGlyValIleuGlyPhePheSerThrIyAspGlnHlsAlaIyGly 191
Qy 595 AACTGGGCTTCAAGACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
Db 192 AsnTrpIyIyHlsLeuAspArgIleValAlaIleuAspGlyTrpIleGlnIleAsnIleAlaHls 211
Qy 655 TTGGTGGGAGCCCGAAGCTGTGACCATGTTTGGCAGTCCGCGGAGGCATTAAGTGT 714
Db 212 PheGlyGlyIyAspProAspArgValThrIlePheGlyGlyIySerAlaIyIyThrSerVal 231
Qy 715 TCAAGCTTATACGTGCTCCATGCGCAAGGCTTATTCACAAAGCCATCAAGAGAT 774
Db 232 SerSerHlsValIValSerProMetSerIyGlyLeuPheHlsGlyAlaIleMetGluSer 251
```


QY	775	GGGGGGGCATCATCCCTTACCTGGAGGCCATGATTGAGGAAGTGAAGACTGGAC	834
Db	252	GIYVAlAlAlAluLeuProTyrLeuIeThraBpThrSerGluMetValSer-----Thr	269
QY	835	GTGGTGTGCAATTTCTGTGTAAACATCGTCAGACTGTGAGCGCCTGTGAGTGCCTG	894
Db	270	ThrValAlAlAluLeuSerGlyCyseGluAlaMetBpSerGluAlaLeuValArgCyseLeu	289
QY	895	AGACAAACCCCTCCAGAGACTGTGACCCCTCAGCCCAAGAAACAAAGTCTTTCACTGCA	954
Db	290	ArgGlyLeuSerGluAlaGluIleLeuAlaIleAluLeuValGluMetIleProAla	309
QY	955	GTGGTGTATGTGGCTTTCTTTCTCTTAATGAGCCTCTAGATCTAATGTCTCAGAAAGCATTT	1014
Db	310	ValValAlaBpGlyGluBpPheProArgH1BpGlyGluLeuAlaSerGluBpPhe	329
QY	1015	AAAGCAATTCCTTCATCATGCGAGTCAATAAACCGAGTGTGGCTCTCTGCTGCCTATG	1074
Db	330	HisProValProSerIleIleIleGlyAlaBpBpBpGluBpIleGlyTyrThrIlePro---	348
QY	1075	AAGAGGCTCTGAGATCTCTAGTGGCTTCAACAAAGTCCCTTGCCCTCATCTGTATCAAA	1134
Db	349	-----ValValMetGlySerAlaGlnThrIleLeuGluIleThraArgGlu	363
QY	1135	AACATC-----CTGCATATCCGCCCTCAGATATTG	1166
Db	364	AsnLeuGlnAlaValLeuLysAsnThrThraIaGlnLeuMetLeuProProGluYsSer	383
QY	1165	CACCTTGAGCTTAATGAAATCTCTCATATACAAGACCTCCCTGACTGAATGCAAGACAGT	1224
Db	384	AspLeuLeuMetGluGluTyrMetGlyAspThrGluBpAlaGlnThrLeuGlnIleGln	403
QY	1225	CTTCTGACCTTGAGATGTGTCTTTGTGTGCTCCCTGCACTGATCAGACCTGATAT	1284
Db	404	PheThrGluMetMetGluAspPheMetPheValIleProAlaLeuGlnValAlaIleThre	423
QY	1285	CACAGAGATGCTGGTGCACCTGTCTACTTCTATAGATTTCGGACACCGGCTCAGTGGCTTT	1344
Db	424	GlnArg--SerH1AlaSerValTyrPheTyrGluBpGlnH1GlnIleAlaSerLeu	442
QY	1345	GAAAGACAAAGCGCGCTTTGTCAACCGGACCGCATGCTGAATGATCGGCTTGCTGTCT	1404
Db	443	LysAspValAlaGpProThnH1ValLysAlaBpH1ValaBpGluIleProPheValPhe	462
QY	1405	GCTGGTGCCTC-----CTGAAGGGGACATTTGTTATGTTCAAGGACCAAGGAGAG	1458
Db	463	GlyTyrPhePheIleThrAspMetLysLeuAspPhe-----ThrGluGly	476
QY	1459	GAGAAATTACTGAGCCCGGAAGATGATGAATATCTGGGCTACCTTGTCTGAACCGGAAAT	1518
Db	477	GluLysLeuLeuSerAlaArgMetMetLysTyrTrrAlaAspPheAlaArgH1GlyLys	496
QY	1519	CTTAATGGGAACGACTGTCTGTGGGCAAGCTTAATCTGACTGACAGACAGTCCAG	1578
Db	497	ProAsnSerGluGlyLeuProTyrTrrProValMetAspH1AspGluGlnTyrLeuGln	516
QY	1579	CTGCACTTGAACATGAGCCTCGGACAGAGACTCAAGAAACCGCGGCTGATTTTGGAC	1638
Db	517	LeuAspThrGlnProAlaValAlaGlyArgAlaLeuLysSerAlaGlyLeuGlnPheThr	536
QY	1639	AGACACATCCCC-----CTGATCTGTCTGTCTCCGACATGCTCCACAGTCTCTT	1698
Db	537	LysThrLeuSerGlnLysIleGlnGluLeuArgAlaSerGlnAspLysH1ThrGluLeu	556
RESULT 14			
Q6GM54	ID	PRELIMINARY;	PRT; 568 AA.
Q6GM54;	AC	Q6GM54;	
DT	05-JUL-2004	(TRIMBLrel. 27, Created)	
DT	05-JUL-2004	(TRIMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TRIMBLrel. 27, Last annotation update)	
DE	LOC443703	protein (Fragment).	
GN	Name=LOC443703;		

OS Xenopus laevis (African clawed frog).
OC Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Butelosteomi;
NC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney.
EX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Scheer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haefl F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ushed T.B., Toshiyuki S., Carrinetti P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smalins D.E., Schermer A., Schein J.B.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney.
EX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.N., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL, BC074230; AAH74230.1; --
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR0002018; CarpesteraseB.
DR InterPro: IPR000379; Ser_estr.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KM Hydrolyase.
FT NON TER
SQ SEQUENCE 568 AA; 63032 MW; 9336D0B6B53931A1 CRC64;

Alignment Scores:
Pred. No.: 2,64e-85 Length: 568
Score: 1263.50 Matches: 258
Percent Similarity: 64.49% Conservative: 87
Best Local Similarity: 48.22% Mismatches: 169
Query Match: 39.86% Indels: 21
DB: Gaps: 8

US-10-023-515-3 (1-1746) x QGGM54 (1-568)

OY 82 GGAAAACTGGGCTTTCGTGCCTGAAGGGCAGACAGAGAACCAGAGCTGGAGTGATTCAG 1411
Db 28 GYIYLtGYrGlyThncInuSpAlarGrProleuWthrTrhAnenryrGLyGlnLeuLeu 47
OY 142 GCGACGACGACTGCTGTGCTGGAAAGCCCTGTGCTGTAACGCTTTCTCGGAGTCCCC 201
Db 48 GilyLyrrThValGlyVAlaLySGluThrSPrArgLeuIleuValPheMetGlyValPro 67


```

Oy 58 CCCCTGTTGGACACAGACAGTGGGAAAACTGGGCTTGTCTGAAGGGCCACAGAG 117
Db 20 ---LeuLeuIleHisValGln---GlyGlnAspSerProGluSer---SerProIleArg 36
Oy 118 AACACGAGCTGGAGTGGATTCAGGGCAAGCAAGTCACTGTCTGGAGCCCTTGCT 177
Db 37 ThrTrhTrhTrhGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 56
Oy 178 GTGAACGTTCTCTGGAGAGTCCCTTTGTCTGCTCCCGCTGGAGTCCCTGGATTAC 237
Db 57 ValHisTrhTrhPheLeuGlyIleProPheAlaLysProValGlyProLeuTrhPheAla 76
Oy 238 AACCCGAGCTGCATCGCCCTGGAGTAACTGCGAGAGACCACTCCCTACCTCAATTG 297
Db 77 ProProGluProProGluProGluProTrpSerGlyValArgAspAlaTrhSerGlnProAlaMet 96
Oy 298 TGCCTCCAGAACTCAGAGTGGCTG-----CTCTTAAGATCAACATGCTCAAG 345
Db 97 CysLeuGlnIleAsnLeuAspIleLeuAspGluValGlyLeuLeuAspMetIle--- 115
Oy 346 GTGCATTACCCGAATTGGAGATGCAGAGATGCTCTTACCTTGAACATCTATCGCT 405
Db 116 ---LeuSerSerIleSerMetSerGluAspCysLeuTyTrLeuAsnValTyTrAlaPro 133
Oy 406 GCCCAGCCGATACAGGCTCCAAAGCTCCCGCTTGTGTGGTCCAGAGAGTCCCTTC 465
Db 134 AlaHisAlaAspGluGlnGlnSerAsnLeuProValMetValTrpIleHisGlyIleValLeu 153
Oy 466 AAGACTGCTCAGCTCCATCTTGTATGGTGGCTCCGCTGGCTGCTATGAGAGACGTGCTG 525
Db 154 ValValGlyMetAlaSerMetCysTrhAspGlySerLeuLeuTrhValAsnGluAspLeuVal 173
Oy 526 GTTGTGCTGCTCAGTACCGGCTAGAGATATTGGTTCTTACCAACATGGATCAGCAT 585
Db 174 ValValTrhTrhIleGlnTyArgLeuGlyValLeuGlyPhePheSerTrhGlyAspGlnHis 193
Oy 586 GCTCCGGGGAACTGGGCTTCAAGACCAAGGCTGCTGCTGGGCTCCAGAAAGAC 645
Db 194 AlaArgGlyAsnTrpGlyTyTrLeuAspGlnValAlaAlaLeuArgTrpValGlnGlnHis 213
Oy 646 ATCGAGTCTTCTGGTGGAGCCCAAGCTCTGTGACCATCTTGTGGAGAGTCCCGGAGCC 705
Db 214 IleAlaHisPheGlyGlyAsnProAsnArgValTrhIlePheGlyGluSerIleGly 233
Oy 706 ATAACTTTTCTAGTCTTAACTGTCTCCCATGSCCAAAAGCTTATTCACAAAGCATC 765
Db 234 ThrSerValSerSerHisValIleSerProMetSerGlnGlyLeuPheHisGlyAlaIle 253
Oy 766 ATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGCCCAAGATTAAGAAAGTGA 825
Db 254 MetGluSerGlyValAlaLeuLeuProAspLeuIleSer-----GluTrhSerGlu 270
Oy 826 GACCTGCAG---GTGGTTGSCATTTCTGTGTAAACAATGCGTCAAGACTGTAGGCCCTG 882
Db 271 ThrValSerTrhTrhValAlaLysLeuSerGlyCysGluAlaMetAspSerGlnAlaLeu 290
Oy 883 CTGAGAGTGCCTGAGGACAAACCTTCAAGAGCTGTGACCCCTCAAGCCAAACAAAG 942
Db 291 ValArgCysLeuArgAlaLysSerGlyAlaGluIleLeuValIleAsnLysValPheLys 310
Oy 943 TCTTTCACTCGAGTGGTGAATGTGCTTCTTCTTAATGAGCTCTGAATCTATGTCT 1002
Db 311 MetIleProAlaValValAspIleGluProArgHisProLysGluLeuLeuAla 330
Oy 1003 CAGAAAGCATTTAAAGCAATCTCTTCAATCATGAGTCAATACCAACGAGTGGCTTC 1062
Db 331 SerGluAspPheHisProValProSerIleIleGlyValAsnTrhAspGluTyTrCysCys 350
Oy 1063 CTGCTGCTTANG-----AAGAGAGCTCTGAGATCTCTCACT 1098
Db 351 ThrIleProMetValMetGlyTrhAlaGlnIleIleLysGlu-----LeuSer 366

```

```

Oy 1099 GACTTCAAACAAGTCCCTTGGCCCTCCATCTGATACAAAACATCTGCATCCCGCTCAG 1158
Db 367 ArgGluAsnLeuGlnAlaValLeuLysAspTrhAlaAlaGlnMetLeuProProGlu 386
Oy 1159 TATTTCACCTTTGGCTTAAGATTAATCTTCCATGACCAAGCACTCCCTACCTGAATACTCGA 1218
Db 387 CysGlyAspLeuLeuMetGluGlnTyTrMetGlyAsnTrhAspAspSerGlnTrhLeuGln 406
Oy 1219 GACAGCTCTTGCAGCTTGCCTGAGAGATGTCTTCTTGTGGTCCCTGCACTGATCACACT 1278
Db 407 IleGlnTyTrhGluMetMetGlyAspPheLeuPheValIleProAlaLeuGlnValAla 426
Oy 1279 CGATATCAACAGAGATGCTGTGACCTGTCTTACTTATGAGTTTGGCACCGGCTCAG 1338
Db 427 HisPheGlnArg---SerHisAlaProValTyTrPheTyTrGluPheGlnHisAlaProSer 445
Oy 1339 TGCCTTGAACACAGAACCGGCTTTTGTCAAGGCCCAACGCTGATGAAGTCCGCTTT 1398
Db 446 TyTrPheLysAsnValArgProProHisValLysAlaAspHisAlaAspGluValProPhe 465
Oy 1399 GTGTCGGTGGTGGCTCTCTGAGAGGGGACATGTGTATGTTCGAAGAGCCAGGAGAG 1458
Db 466 ValPhe---GlySerPhePheSerGlyMetLysLeuAspPhe-----ThrGluGln 481
Oy 1459 GAGAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTTGTCTGAAACCGGGAAT 1518
Db 482 GlnArgLeuLeuSerArgArgMetMetLysTyTrTrpAlaAsnPheAlaArgGlnGlyAsn 501
Oy 1519 CCTAATGGAAACGACCTGTCTGTGGCCAGCTTATTAATCTGACCTGAGCAGTACTCCAG 1578
Db 502 ProAsnSerGluGlyLeuProTyTrTrpProAlaLeuAspHisAspGluGlnTyTrLeuGln 521
Oy 1579 CTGACCTTGAACATGAGCTGTGCTGTGGCCAGCTTATTAATCTGACCTGAGCAGTACTCCAG 1638
Db 522 LeuAspTrhHisProAlaValAspArgAlaLeuLysValAlaArgGluLeuGlnPheTrhTrh 541
Oy 1639 AGCACCATCCCC 1650
Db 542 LysTrhLeuPro 545

```

Search completed: June 15, 2005, 08:53:13
Job time : 183.861 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 07:48:14 ; Search time 26.3868 Seconds

(without alignments)
9878.981 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 3170
1 atgcacagagactactctc.....cttctcttgcctcttga 1746

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p model -DEV=x1h
-Q=/cgn2_1/USPRO_spool/6664091/runat_14062005_133500_9360/app_query.fasta_1.4238
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091 @cgn2_1_154 @runat_14062005_133500_9360 -NCPU=6 -ICPU=3 -NO_MMAP
-LARGEQUERY -NSG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRADs=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	97.1	581	4	US-10-023-515-2
2	1230	38.8	559	4	US-09-595-682B-28
3	1230	38.8	559	4	US-09-949-016-6426
4	1230	38.8	577	4	US-09-949-016-9670
5	1150.5	36.3	565	4	US-09-595-682B-21
6	1138.5	35.9	566	3	US-09-264-737-2
7	1130.5	35.7	543	4	US-09-595-682B-26
8	1130.5	35.4	584	2	US-08-845-295A-2
9	1091.5	34.4	584	3	US-09-140-933-2
10	1091.5	34.4	584	3	US-09-146-661-2
11	1091.5	34.4	584	3	US-09-150-515-2
12	1034.5	32.6	539	3	US-09-264-737-1

13	942.5	29.7	454	3	US-08-446-100-26	Sequence 26, Appl
14	942.5	29.7	454	3	US-08-446-100-28	Sequence 28, Appl
15	942.5	29.7	454	3	US-08-446-100-30	Sequence 30, Appl
16	942.5	29.7	454	3	US-08-446-100-31	Sequence 31, Appl
17	940.5	29.7	454	3	US-08-446-100-27	Sequence 27, Appl
18	940.5	29.7	454	3	US-08-446-100-29	Sequence 29, Appl
19	871	27.5	574	4	US-10-023-515-4	Sequence 4, Appl
20	793	25.0	933	4	US-09-949-016-8386	Sequence 8386, Ap
21	786	24.8	823	4	US-09-949-016-6888	Sequence 6888, Ap
22	782	24.7	953	4	US-09-949-016-8387	Sequence 8387, Ap
23	777.5	24.5	843	4	US-09-491-356C-20	Sequence 20, Appl
24	777	24.5	575	1	US-08-348-920-1	Sequence 1, Appl
25	775	24.4	823	4	US-09-491-356C-23	Sequence 23, Appl
26	774	24.4	575	1	US-08-348-920-2	Sequence 2, Appl
27	772	24.4	848	4	US-09-491-356C-22	Sequence 22, Appl
28	769	24.3	836	4	US-09-491-356C-21	Sequence 21, Appl
29	765.5	24.1	614	3	US-08-446-100-25	Sequence 25, Appl
30	764.5	24.1	614	1	US-07-732-962A-2	Sequence 2, Appl
31	764.5	24.1	614	2	US-08-370-156-2	Sequence 2, Appl
32	764.5	24.1	614	3	US-08-446-100-19	Sequence 19, Appl
33	764.5	24.1	614	3	US-08-446-100-21	Sequence 21, Appl
34	764.5	24.1	614	3	US-08-814-095-2	Sequence 2, Appl
35	764.5	24.1	614	5	PCT-US92-06106-2	Sequence 2, Appl
36	764.5	24.1	645	4	US-09-949-016-7063	Sequence 7063, Ap
37	764.5	24.1	645	4	US-09-949-016-7064	Sequence 7064, Ap
38	762.5	24.1	614	3	US-08-446-100-20	Sequence 20, Appl
39	761.5	24.0	614	3	US-08-446-100-23	Sequence 23, Appl
40	760.5	24.0	614	3	US-08-446-100-22	Sequence 22, Appl
41	759.5	24.0	600	2	US-08-370-156-4	Sequence 4, Appl
42	759.5	24.0	600	3	US-08-814-095-4	Sequence 4, Appl
43	755.5	23.8	617	2	US-08-975-084-1	Sequence 1, Appl
44	755.5	23.8	617	2	US-08-370-156-6	Sequence 6, Appl
45	755.5	23.8	617	3	US-08-814-095-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010. A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:
Pred. No.: 2,866-294
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 97.13%
DB: 4
Gaps: 0
US-10-023-515-3 (1-1746) x US-10-023-515-2 (1-581)
QY 1 ATGCACAGGAGACTTACATCTGCTTCACATGGTGTCTTTCCTGATTCGAGCCC 60
|||||

Db 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
QY 61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
Db 21 LeuLeuGlnYHSAAGlnTrpGlySerThrGlyProSerAlaGlnGlyProGlnTrpGln 40
QY 121 ACCAGGCTGGAGTATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGGCTGTG 180
Db 41 ThrArgLeuGlnGlyTrpIleGlnGlySerGlnValThrValLeuGlySerProValProVal 60
QY 181 AACGTTTCTCTGGAGTCCCTTTGCTGTCCCGCTGGAGATCCCTGGATTAACGAAC 240
Db 61 AsnValPheLeuGlyValProPheAlaAlaProPheLeuGlySerLeuArgPheThrAsn 80
QY 241 CCGAGCTGGATGGCCCTGGGATTAACCTTGGCAAGAACCACTCCTACCTTAATTTTGGC 300
Db 81 ProGlnProAlaSerProTrpAsnLeuArgGlnValAlaThrSerTrpProAsnLeuGly 100
QY 301 CTCCAGAACTGAGAGTGGCTGCTTGAATCAACACATGCTCAAGTGCATTACCCGAAA 360
Db 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHISMetLeuLeuValHisTrpProlys 120
QY 361 TTCGAGTGTCAAGAACTGCTCTTAACTGAACATCTATGCGCTGCCACGCGCATACA 420
Db 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
QY 421 GGCTCCAAAGCTCCCGCTGGGTGTGTTCCAGGAGGAGCTTCAAGACTGGCTCAGCC 480
Db 141 GlySerIleLeuProValLeuValTrpPheProGlyGlyAlaPheLeuTrpHisSerAla 160
QY 481 TCCATCTTTGATGAGTCCGCCCTGGCTGCTATGAGAGAGTGTGGTGTGGCTCCAG 540
Db 161 SerIlePheAspGlySerAlaLeuAlaIleTrpGlnAspValLeuValValGln 180
QY 541 TACCGGCTAGAAATTTGGTTTCTTACACACATGGGATCAGCATGCTCCGGGAACTGG 600
Db 181 TyrTrpLeuGlyIlePheGlyPhePheThrTrpAspGlnHISAlaProGlyAsnTrp 200
QY 601 GCCTTCAAGACAGAGTGGCTGTCTGCTGGGTCCAGAAACATCGATTCCTTCGGT 660
Db 201 AlaPheIleAspGlnValAlaAlaLeuSerTrpValGlnIleAsnIleGlnPhePheGly 220
QY 661 GGGGACCCAGCTCTGTGACCATCTTTGGAGTCCGGGAGCCATTAAGTGTTCAGT 720
Db 221 GlysAspProSerSerValThrIlePheGlyGlnSerAlaIleSerValSerSer 240
QY 721 CTTATACCTGTCTCCAGGCAAGGCTTATTCACAAAGCCATCAGTGAAGTGGGTG 780
Db 241 LeuIleLeuSerProMetAlaIleGlyLeuPheHisIleValAlaIleMetGlnSerGlyVal 260
QY 781 GCCATCATCCCTTACCTGAGGAGCCCATGATTATGAGAGTGAAGAGCTCAGGTGGT 840
Db 261 AlaIleIleProGlyTrpLeuGlnAlaHisAspTrpGlyIleSerGlnAspLeuGlnValVal 280
QY 841 GCACATTTCTGTGTAAACATGCGTCACTTGAAGCCCTGTGAGGTGCTTGAAGACA 900
Db 281 AlaHisPheCysGlyAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
QY 901 AAAACCCCAAGAGCTGCTGACCCCTCAGCCAGAAACAAAGTCTTTACCTCAGTGGT 960
Db 301 LysProSerIleGlyIleLeuLeuThrLeuSerGlnIleTrpLysSerPheThrArgValVal 320
QY 961 GATGATCTTTCTTCTTCTATGAGCTTGAATGATGATGATGATGATGATGATGATGAT 1020
Db 321 AspGlyIleAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnIleValPheIleVal 340
QY 1021 ATTCTTCATCATGAGAGTCAATPACCAAGAGTGTGGCTTCTGCTGCTTATGAAGAG 1080
Db 341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLeuGln 360
QY 1081 GCTCCTGAGATCCTCAGTGGGTCCAAACAAGTCCCTGGCCCTCAATGATCAAAAACATC 1140
Db 361 AlaProIleIleLeuSerGlySerAsnLysSerLeuAlaLeuHisIleGlnAsnIle 380

QY 1141 CTGCACATCCCGCTCAGTATTTTGACCTTGGCTAATGAATCTTCATGACAGACAC 1200
Db 381 LeuHisIleProProGlnTrpLeuHisIleValAlaAsnGlnTrpPheHisAspIleHis 400
QY 1201 TCCCTGATGAAATCCGAGACAGTCTTCTGAGACTTGTGAGATGTGTCTTTGTGCTC 1260
Db 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlnIleAspValPhePheValVal 420
QY 1261 CCTGACATGATCAAGCTCGATATTCACAGAGATGTGTGACCTGTCTTACTTCTATGAG 1320
Db 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGln 440
QY 1321 TTTGCGACCGGCTCAGTGTCTTGAAGACAGAGCCGGCTTTGTCAAAAGCCGACAC 1380
Db 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
QY 1381 GCTGATGAAGTCCGCTTGTGTGCTGTGCTCTTCTGAAGGGGACATTGTTATGTTTC 1440
Db 461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
QY 1441 GAAGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGAAATATCGGCTTACC 1500
Db 481 GlnGlyAlaThrGlnGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpAlaThr 500
QY 1501 TTTGCTGAAACCGGAAATCCTAATGGGAAAGACCTGTCTGTGGCCAGCTTATATCTG 1560
Db 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
QY 1561 ACTGAGACATGACTTCCAGCTGGAATTAAGACCTTGGACAGAGACTCAAGAAACCG 1620
Db 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGlnPro 540
QY 1621 CGGGTGAATTTTGGACACAGACATCCCTGATTCCTGTGGCTCCGATATGCTCCAC 1680
Db 541 ArgValAspPheTrpTrpSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
QY 1681 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCAGCTTCTTCTTCTTGTGCT 1740
Db 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
Db 581 Pro 581

RESULT 2
US-09-595-682B-28
Sequence 28, Application US/09595682B
Patent No. 6800483
GENERAL INFORMATION:
APPLICANT: Danks, Mary K.
APPLICANT: Potter, Philip M.
TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
FILE REFERENCE: SJ-0005
CURRENT APPLICATION NUMBER: US/09/595,682B
CURRENT FILING DATE: 2000-01-16
PRIOR APPLICATION NUMBER: 60/075,258
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: PCT/US99/03171
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 559
TYPE: PRT
ORGANISM: Homo sapiens
US-09-595-682B-28

Alignment Scores:
Pred. No.: 4,44e-112 Length: 559
Score: 1230.00 Matches: 248

Percent Similarity:	63.84%	Conservative:	91
Best Local Similarity:	46.70%	Mismatches:	16
Query Match:	38.80%	Indels:	24
DB:	4	Gaps:	8

[illegible][illegible]

```

DB: 4 Gaps: 8
US-10-023-515-3 (1-1746) x US-09-949-016-6426 (1-559)
QY 91 GGGCTTGTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGAA 150
Db 26 G1ylnApsSerAlaSerPro1LeaG1Thr1Thn1sThG1yAlaValLeu1d1ySerLeu 45
QY 151 GTCACTGTGCTGGAGAGCCCTGCTGCTGAAAGTTCCTGCGAGATCCCTTGTGCT 210
Db 46 ValnIsvallyvG1yAlaAenAlaG1yAlaG1nThrPheLeuG1y1LeProPheAla1y 65
QY 211 CCCCCGTGGAGTCCCTGGATTTAGAACCCGAGCCCTGATCCGCTGGAGTAATGTG 270
Db 66 ProProLeuG1yProLeuA1yPheAlaProProG1uProProG1uSer1ySerG1yAl 85
QY 271 CGAAGAGCACTCCTTACCTTAATTGCTGCTCCAGAAC-----TCAGAG 315
Db 86 ArgAerG1yThr1Thn1sProAlaMetCysLeuG1nAerLeuThrAlaValG1uSerG1u 105
QY 316 TGGGCTGCTTAGATCAACACATGCTCAAGGTGATTCACCGAAATGCGAGTGCAGAA 375
Db 106 PheLeu-----SerG1nAerAenMet1ThrPheProSerAerSerMetSerG1n 121
QY 376 GACTGCTCTATCCTGACATATATGCTGCTGCTCCAGCCGATACAGGCTCCAGCTCC 435
Db 122 AaPcyAer1yThrLeuSer1LeTyThrProAla1sSerTh1sG1y1SerAla1uPro 141
QY 436 GTCTTGTGTGTGCTCCAGAGAGTCCCTTCAAGACTGCTGACCTCCATCTTTATGG 495
Db 142 ValMetVal1Thr1Le1n1sG1yAlaLeuVal1PheG1yMetAlaSerLeu1yThrAer1y 161
QY 496 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
Db 162 SerMetLeuAlaAlaLeuG1uAenVal1Val1Al1e1nG1nThy1ArgLeuG1yAl 181
QY 556 TTTGGTTTCTTACCAACATGGAGATGAGATGCTCCGGGAACTGGGCTTCAAGAGCAG 615
Db 182 LeuG1yPhePheSer1ThrG1yAer1yTh1sAla1ThrG1yAen1ThrG1y1LeuAerG1n 201
QY 616 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Db 202 ValAlaAlaLeuA1yThrVal1G1n1d1nAen1Th1sAla1n1sPheG1y1AerProAer1y 221
QY 676 GTGACATCTTTGGAGAGTCCGCGGAGCCATTAAGTCTTCTAGCTTATCTGCTGCC 735
Db 222 ValThr1LePheG1yG1uSerAlaG1y1yThrSerVal1SerSerLeuVal1Val1SerPro 241
QY 736 ATGGCCAAAGCTTATTCACAAAGCCATGAGAGTGGGAGTGGGCAATCCCTTAC 795
Db 242 1LeSerG1nG1yLeuPhe1n1sG1yAla1LeMetG1uSerG1yVal1AlaLeuProG1y 261
QY 796 CTGAGAGCCCATGATTTAGAGAAGTGAGACCTG-----CAGTGTGGACATTTTC 849
Db 262 Leu1LeAla-----SerSerAlaAerVal11LeSerTh1Val1ValAlaAenLeu 277
QY 850 TGTGTAACAATGGGCTGAGACTCTGAGGCTGCTGAGAGCTGAGAGCAAAACCTCC 909
Db 278 SerAlaCysAerG1nValAerSerG1uAlaLeuVal1G1yCysLeuAerG1y1ySer1y 297
QY 910 AAGAGCTGCTGACCTCAGCCACAGAAAACAAGTCTTCTACCTCAGAGTGGTGTGCT 969
Db 298 G1uG1u1LeLeuAla1LeAen1yProPhe1yMet1LeProG1yVal1ValAerG1yAl 317
QY 970 TTCTTTCTATGAGCTCTGATCTATTTGTCTAGAAAGATTTAAAGCAATTCCTCC 1029
Db 318 PheLeuProAerG1n1sProG1nG1uLeuLeuAlaSerAlaAerPheG1nProVal1ProSer 337
QY 1030 ATCATCGAGCATTAACAACAAGTGTGCTTCTGCTGCT-----ATGAAG 1077
Db 338 1LeVal1G1yAlaAenAenAenG1uPheG1yThrPhe1u1ePro1yVal1MetAerG1le1y 357
QY 1078 GAGGCTCTGAGATCTCTAGTGGCTCAACAAGTCCCTTGCCTTCATCTGATACAAAAC 1137

```

```

Db 358 AaPThrG1n1yG1uMetAerAerG1uAlaSerG1nAla1LeuG1n1yMetLeuThr 377
QY 1138 ATCTGACATCCCGGCTCAGTATTTGCACTTGGGCTAATGAATCTTCAAGCAAG 1197
Db 378 LeuLeuMetLeuProProThrPheG1yAerLeuAerG1uG1u1y1LeG1yAerAen 397
QY 1198 CACTCCCTGAGTGAATCCGAGACAGTCTTCTGAGCTTGTGAGATGATGTTCTTGTG 1257
Db 398 G1yAerProG1nThrLeuG1nAlaG1nPheG1n1yMetMetAlaAerSerMetPheVal 417
QY 1258 GTCCCTGACTGTATCAGCTGATATACAGAGATGCTGTGCACTGTCTTCTAT 1317
Db 418 1LeProAlaLeuG1nValAla1sPhe---G1nCySerAerAlaProVal1yTher1y 436
QY 1318 GAGTTTGGGACCGGCTCAGTGTGTTGAAGACAGAACCGGCTTTGTCAAGCCGAC 1377
Db 437 G1uPheG1n1sG1nProSer1ThrLeu1yAen1LeAerGProPro1sMetLe1yAlaAer 456
QY 1378 CAGCTGATGAAGTCCGCTTGTGTGCTGCTGCTTCTGAAAGGGGACATTGTATG 1437
Db 457 HisG1yAerG1uLeuProPheVal1Phe---AerSerPhePheG1yG1yAenThy1Le1y 475
QY 1438 TTGAAAGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGAATTAATGGCT 1497
Db 476 Phe-----ThrG1uG1uG1uG1uG1u1nLeuSerAerAerG1yMetMetLe1yThrAla 492
QY 1498 ACCTTGTGCTGACCCGGGAATCTTAATGGAGAACACTGTCTGTGCGGCACTTAAT 1557
Db 493 AaenPheAlaAerAenG1yAerProAenG1yG1uG1yLeuPro1sThrProLeuAerAer 512
QY 1558 CTGACTGAGCAGTACTCTCAGCTGAGCTTGAACATGAGACCTCGAGAGACTCAAGAA 1617
Db 513 G1nG1uG1uG1nThy1LeuG1n1yAenLeuAenG1nProAlaVal1G1yAlaLeu1yAla 532
QY 1618 CCGGCGGTGATTTTGGACCAACCAATCCCC 1650
Db 533 HisAerG1uG1nPheThr1y1yAlaLeuPro 543

RESULT 4
US-09-949-016-9670
; Sequence 9670, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OR INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PASTESQ for Windows Version 4.0
; SEQ ID NO 9670
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9670

Alignment Scores:
Pred. No.: 4,53e-112 Length: 577
Percent: 1230.00 Matches: 248
Score Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 38.80% Indels: 24
DB: 4 Gaps: 8
US-10-023-515-3 (1-1746) x US-09-949-016-9670 (1-577)

```



```

OY 91 GGGGCTTGTGGAAGGGCAACAGAGAAACACAGAGCTGGATGTTCAGGCAAGCA 150
DB 44 G1ylnaPserAlaSerPro1LeaGhrThrHnIsthrGlyInValaLeuG1Serleu 63
OY 151 GTCACTGTGTGGGAAGCCCTGTGCTGTGAACSTGTTCCTGAGAGTCCCTTGTCTGT 210
DB 64 ValHnValySg1yAlaHnAlaG1yAlaG1nThrPheLeuG1y1LeProHnAlaLys 83
OY 211 CCCCCGTGGGATCCCTGGATTTACGAACCCGACGCTGCATGCCCTGGATTAATTG 270
DB 84 ProProLeuG1yProLeuAlaGhrPheAlaProProGluProProGluSerTrpSerG1yAl 103
OY 271 CGAGAAGCACCTCTACCTATTTGTGCTCCAGAAC-----TCAGAG 315
DB 104 ArgHnArg1yThrHnHnProAlaMetCySeuGlnHnSerLeuThAlaValG1uSerGln 123
OY 316 TGGCTGTCTTAGATCAACACATGCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAG 375
DB 124 Pheleu-----SerGlnPheHnMetThrPheProSerArgSerMetSerGln 139
OY 376 GACGTGCTCTACCTGAACATCTATGCGCTGCCACGCGGATACAGCTCCAAAGCTCCC 435
DB 140 AspCySeuLeuTyLeuSer1LeTyThrProAlaHnSerHnSg1ySerHnLeuPro 159
OY 436 GTCCTGTGTGGTCCCAAGAGGTGCTTCAAGAGCTGCTGACCTTCATGGATGGG 495
DB 160 ValMetVal1Trp1LeHnSg1yAlaLeuVal1PheG1yMetAlaSerLeuTyArgG1y 179
OY 496 TCCGCGCTGTGCTATAGAGACGCTGTGCTGTGTGTCGTCACATACCGCTAGAAATA 555
DB 180 SerHnLeuAlaAlaLeuGlnHnVal1Val1Val1Le1LeG1nTyArgLeuG1yAl 199
OY 556 TTTGGTTTCTTCAACCATGGGATGAGATGCTCCGGGAACTGGGCGCTTCAAGACAG 615
DB 200 LeuG1yPhePheSerThrG1yAspTyHn1Sa1aThrG1yAsnTrpG1yTyTyLeuAspGln 219
OY 616 GTGGCTGTCTGTCTCTGGTCCAGAAACATCGAGTTCTTCGGTGGGAGCCCGAGCTCT 675
DB 220 ValAlaAlaLeuArgTyValGlnHnHnAlaHn1PheG1yG1yAsnProAspArg 239
OY 676 GTGACATCTTTGGGAGAGTCCGCGGAGCCATTAAGTTTCTTATGCTTATCTGCTCC 735
DB 240 ValHnThr1LePheG1yG1uSerAlaG1yG1yThrSerValSerSerLeuValValSerPro 259
OY 736 ATGGCAAGGCTTATTCACAAACCATGATGAGAGTGGGCTGACCATCCCTTAC 795
DB 260 1LeSerGlnG1yLeuPheHnSg1yAla1LeMetG1uSerG1yValAlaLeuLeuProG1y 279
OY 796 CTGAGAGCCCATGATTAAGAAAGTGAAGCTG-----CAGGTGGTTCACATTTTC 849
DB 280 Leu1LeAla-----SerSerAlaAspVal1LeSerThrValValAlaHnSerLeu 295
OY 850 TGTGGTAAATGCGTCAAGCTCTGAGGCGCTGTGAGGTGCTGAGGACAAACCTTCC 909
DB 296 SerAlaCySeuArgGlnValaAspSerGlnAlaLeuValG1yCySeuLeuArgG1yLysSerLys 315
OY 910 AAGAGCTGTGAGCCCTCAAGCAAAACAAAGCTTTTCACTGAGTGGTGTGAGTGTCT 969
DB 316 GluGlu1LeLeuAla1LeAsnTyAspProPheTyMet1LeProG1yValValaAspG1yAl 335
OY 970 TTCTTTCTTAATGAGCTCTAGATCTATTTGTCAAAAAGCTTTAAAGCAATTCCTTCC 1029
DB 336 PheLeuProAlaGhr1AspProGlnG1uLeuLeuAlaSerAlaAspPheGlnProVal1ProSer 355
OY 1030 ATCATCGAGTCAATTAACCAACAGATGTGGCTTCTGTGCTCT-----ATGAAG 1077
DB 356 1LeValG1yValaAsnHnSerGln1uPheG1yTrpLeu1LeProLysValaMetArgG1eTy 375
OY 1078 GAGGCTCTGAGATCTCTAGTGGCTTCAACAAAGCTTCCTTGCATCTGATTAACAAAC 1137
DB 376 AspThrGlnTyG1uMetAspArgG1uAlaSerGlnAla1LeuGlnTyMetLeuThr 395

```

```

OY 1138 ATCTGCATATCCGCGCTCAGTATTTGCAACCTTGTGCTAATGAATTACTGCATGACAAG 1197
DB 396 LeuLeuMetLeuProProThrPheG1yAspLeuLeuArgGlnG1uTy1LeG1yAspHn 415
OY 1198 CACTCCCTGCATGAAATCCGAGACAGCTTCTTGCAGCTTGTGGAGATGCTGTTTGGG 1257
DB 416 G1yAspProGlnThrLeuGlnAlaGlnPheGlnGlnMetMetAlaAspSerMetPheVal 435
OY 1258 GTCGCGACATGATACAGACGCTGCATATACAGAGATGCGGTGACCTGTCTACTTCTAT 1317
DB 436 1LeProAlaLeuGlnValAlaHnPhe---GlnCySerArgAlaAspProAlaTyTrpHnTy 454
OY 1318 GAGTTTCGGAACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAGCCGAC 1377
DB 455 GlnPheGlnHnSg1nProSerTrpLeuTyAsn1LeArgProProHn1MetCySAlaAsp 474
OY 1378 CACGCTGATGAAGTCCGCTTGTGTGTCGATGCTCTTCAAGGGGACATTTGATG 1437
DB 475 HnSg1yAspG1uLeuProPheValPhe--ArgSerPhePheG1yG1yAsnTyX1LeTy 493
OY 1438 TTGCAAGAGCAACGAGAGAGAGAAATTACTGAGCCGGAAGATGATGAATTACTGGGCT 1497
DB 494 Phe-----ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 510
OY 1498 ACCTTGTCTGGAACCGGGAATCTTATGGAACGACCTGTCTGTGCGCACTTATAT 1557
DB 511 AsnPheAlaArgAsnG1yAsnProAsnG1yGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 530
OY 1558 CTGATGACAGATACCTTCAGCTGACTTGAACATGAGCTTGGACAGAGACTCAAGAA 1617
DB 531 GlnGlnGlnGlnTyLeuGlnGlnLeuAsnGlnGlnProAlaValG1yArgAlaLeuTySAla 550
OY 1618 CCGCGGTGAGATTTTGGACAGCAACCATCCCC 1650
DB 551 HnArgLeuGlnPheTrpTyAlaLeuPro 561

RESULT 5
US-09-595-682B-21
/ Sequence 21, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danks, Mary K.
/ APPLICANT: Potter, Philip M.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE OF INVENTION: Tumor Cells
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR FILING DATE: 2000-01-16
/ PRIOR APPLICATION NUMBER: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR APPLICATION NUMBER: PCT/US99/03171
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 21
/ LENGTH: 565
/ TYPE: PRT
/ ORGANISM: Oryctolagus cuniculus
US-09-595-682B-21

Alignment Scores:
Pred. No.: 3,05e-104 Length: 565
Score: 1150.50 Matches: 252
Percent Similarity: 56.97% Conservative: 75
Best Local Similarity: 43.90% Mismatches: 180
Query Match: 36.29% Indels: 67
DB: 4 Gaps: 13

US-10-023-515-3 (1-1746) x US-09-595-682B-21 (1-565)
OY 34 TGGTGTCTTTCTGATTTCTCCAGCCCTGTGGGACACAGACAGTGGGGAATAAATCGGG 93
DB 34 TGGTGTCTTTCTGATTTCTCCAGCCCTGTGGGACACAGACAGTGGGGAATAAATCGGG 93

```



```

QY 34 TGGTCCTTTTCCGATTCCTCCAGCCCTGTGGAGCAGCAGAGTGGGAAAACTGGG 93
DB 9 Trrleuphe-----LenuaIaCystrAlatrpolyhis----- 20
QY 94 CTTTGTGCTGAAGGCCACAGAGAACACAGGCTGGGATTCAGGGCAAGCAATGC 153
DB 21 ProSerAlaPro---ProValValAspThrValLySgLyValLeuLylYsPheVal 39
QY 154 ACTGTCGGGAGAGCCGTGTCGTGGAACGGTTCCTGGAGTCCCTTGTCTGTCC 213
DB 40 SerLeuGluGlyPheAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 59
QY 214 CCGCTGGATCCCTCGATTCAGAACCCGAGCTGCATCCCTGGGATCACTGCGCA 273
DB 60 ProLeuGlySerLeuAlaArgPheAlaProProGlnProAlaGlnSerTrpSerHisVallys 79
QY 274 GAAGCACCCTCTACCTTAATTGTCCTCCAGAACTCAGAGTGCCTGCTTACATCAA 333
DB 80 AsnThrTrpSerTrpProMetCysSerGlnAspAla-----ValSerGly 95
QY 334 CACATGCTC-----AAGTGCATTAACCG---AAATTCGAGTGC 369
DB 96 HisMetLeuSerGluLeuPheThrAsnArgLySgLyAsnLleProLeuYsPhe----- 113
QY 370 TCAGAGACTGCCTTACCTGAAACATCTATAGCGCTGCCACGCGCATCAGGCTCAAG 429
DB 114 SerGluAspCySLeuTrpLeuAsnLleTrpThrProAlaAspLeuThrLysArgLySg 133
QY 430 CTCCCGCTTGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGCTCAGCCTTCATCTTT 489
DB 134 LeuProValMetValTrpLleHisGlyGlyLeuMetValGlyValAspSerThrTyx 153
QY 490 GATGGTCCTCCCTGCTGCTATAGAGACCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 549
DB 154 AspGlyLeuAlaLeuSerLleHisGlnAsnValValValValThrLleGlnTrpArgLeu 173
QY 550 GGAATATTGTGTTTCTTCACACATGAGTGCATGCTCCGCGGAACTGGGCTTCAAG 609
DB 174 GlyLleTrpGlyPhePheSerThrGlyAspGlnLysSerArgGlyAsnTrpLylHisLeu 193
QY 610 GACCAAGTGGCTGCTCTGTCTCTGCTGCTCCAGAGAACTGAGATTCTTGTGGGAGCC 669
DB 194 AspGlnValArgAlaLeuArgTrpValGlnAspAsnLleAlaAsnPheGlyLysPro 213
QY 670 AGCTGTGACCATTTTGGCGAGTCCCGGGAGCCATAGCTTTCTTATACG 729
DB 214 GlySerValThrLlePheGlyGlnSerAlaGlyGlnSerValSerLleLeuLeuLeu 233
QY 730 TCTCCATGGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGGATGCATATC 789
DB 234 SerProLeuThrLysAsnLeuPheHisArgAlaLleSerGlnSerGlyValAlaLeuLeu 253
QY 790 CTTTACCTGAGGCCCATGATTATGAGAGAGTGAAC-----CTGCAAGTGTGTGCA 843
DB 254 SerSerLeu-----PheArgLysAsnThrLysSerLeuAlaGlnLysLleAla 269
QY 844 CATTTCTGTGTAAACATGCGTCAAGCTTGAGGCCCTGCTGAGTGCCTTGAGACAAA 903
DB 270 LleGlnAlaGlyCySLeuThrThrTrpSerAlaValAlaMetValHisCySLeuArgLys 289
QY 904 CCTCCAGAGAGCTGCTG-----ACCCCTCAGCCAGAA----- 936
DB 290 ThrGlnGlnGlnLeuMetGlnValThrLeuLysMetLysPheMetAlaLeuAspLeuVal 309
QY 937 -----ACAAAGTCTTTCACCTGAGTGTGATGCTTCTTTCTTCT 978
DB 310 GlyAspProLySgLyAsnThrAlaPheLeuThrTrpValLleAspGlyValLeuLeuPro 329
QY 979 AATGAGCTCTGATCTATTTGTCTCAGAAAACATTTAAAGCAATTCCTTCATCGCA 1038
DB 330 LysAlaProAlaGlnLleLeuAlaGlnLysLysTrpAsnMetLeuProLyMetValGly 349
QY 1039 GTCAATTAACAGAGTGTGCTTCTGCTGCTATG----- 1074

```

```

DB 350 IleAsnGlnGlnLysPheGlyTrpLleIleProMetGlnMetLeuGlyTyxProLeuSer 369
QY 1075 -----AAGAGGCTCTCTGATGATCTTCAGTGGCTCAAAAGTCCCTT 1116
DB 370 GlnGlyLysLeuAspGlnLysThrAlaThrGlnLeuLleTrpLysSer----- 385
QY 1117 GCCCTCATCTGATCAAAACATCTCGACATCCCGCTCAGTAAATTGGACCTGTGGCT 1176
DB 386 -----TyxProLleValAsnValSerLysGlnLeuThrProValAlaThr 400
QY 1177 AATGAATCTTCATGACAGACATCCCTGACTGAATCCAGACAGATCTTGGACTTG 1236
DB 401 GlnTyxTrpLeuGlyGlyThrAspAspProValLysLysLysAspLeuPheLeuPheMet 420
QY 1237 CTTGAGATGTCCTTCTTGTGTGCTCCCTGACCTGATCAGACCTCGATACAGAGATGCT 1296
DB 421 LeuAlaAspLeuLeuPheGlyValProSerValaAsnValAlaArgHisArgAspAla 440
QY 1297 GGTGACCTGTCTACTCTTATGATGTTTGGACACCGGCTCAGTGTCTTGAAAGACGAG 1356
DB 441 GlyAlaProThrTrpMetTyxGlnTyxArgTyxArgProSerPheSerSerAspMetArg 460
QY 1357 CCGGCTTTTGTCAAAAGCCGACGACGCTGATGAAGTCCGCTTGTGTGTGTGTGTGT 1416
DB 461 ProLyxThrValLleGlyAspHisGlyAspGlnLlePheSerValLeuGlyAlaProPhe 480
QY 1417 CTGAAGGGGACATTTGTTATGTTTCAGAGAGCCACGAGAGAGAGAACTTACGACCCG 1476
DB 481 LeuLys-----GlnGlyAlaThrGlnGlnGlnLleLysLeuSerLys 494
QY 1477 AAGATGATGAATATCAGGCTACTTGTCTCGAACCGGGAATCTTAATGGAACGACTG 1536
DB 495 MetValMetLysTrpTrpAlaAsnPheAlaArgAsnGlyAsnProAsnGlyGlnGlyLeu 514
QY 1537 TCTGTGGCCAGCTTATATCTGACTGACAGACATCTTCAGCTGCACTTGAACATGACC 1596
DB 515 ProGlnTrpProAlaTyxAspTyxLysGlnGlyTyxLeuGlnLleGlyAlaThrGln 534
QY 1597 CTCGACAGAGCTCAAAAGACCGGGGTGAGATTTTGGACC 1638
DB 535 AlaAlaGlnLysLeuLysAspLySgLyValAlaPheTrpThr 548

RESULT 7
US-09-595-682B-26
; Sequence 26, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danke, Mary K.
; APPLICANT: Poter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 26
; LENGTH: 543
; TYPE: PRN
; ORGANISM: Oryctolagus cuniculus
US-09-595-682B-26

Alignment Scores:
Pred. No.: 2,796-102 Length: 543
Score: 1130.50 Matches: 249
Percent Similarity: 56.84% Conservative: 75
Best Local Similarity: 43.68% Mismatches: 179

```


Db 557 LyseGlygluValAlaPheThrPasnAsePheLeu 568

RESULT 9
US-09-140-933-2
Sequence 2, Application US/09140933
Patent No. 6022719
GENERAL INFORMATION:
APPLICANT: Hubbe, John C.
TITLE OF INVENTION: Enzymatic Process for the Manufacture of
L-Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
L-Ascorbic Acid
TITLE OF INVENTION: 2-Keto-L-Gulonic Acid
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eastman Chemical Company
STREET: P.O. Box 511
CITY: Kingsport
STATE: Tennessee
COUNTRY: USA
ZIP: 37662-5075
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/140,933
FILING DATE: 27-August-98
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,879; 08/845,295
FILING DATE: 17-May-96; 25-April-97
ATTORNEY/AGENT INFORMATION:
NAME: Cheryl J. Tubbach
REGISTRATION NUMBER: 38,346
REFERENCE/DOCKET NUMBER: 70432
TELECOMMUNICATION INFORMATION:
TELEPHONE: 423-229-6189
TELEFAX: 423-229-1239
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-09-140-933-2

Alignment Scores:
Pred. No.: 2,03e-98 Length: 584
Score: 1091.50 Matches: 259
Percent Similarity: 54.39% Conservative: 63
Best Local Similarity: 43.75% Mismatches: 191
Query Match: 34.43% Indels: 79
Gaps: 17

US-10-023-515-3 (1-1746) x US-09-140-933-2 (1-584)

QY 34 TGGTGGCTTTTTCGATTCTCAGCCCTGTGGACACAGACAGTGGGAAAACTGGG 93
Db 2 TrrpeuLeuPrrProValleuThrrSerleuAaserSerAlaThrTrp-----AlaGly 19
QY 94 CCTTGTGTAAGGCGCACAGAGACACAGGCTGGTGGATTCAAGGCAAGCAATTC 153
Db 20 GlnPrrAlaSerPrrProValValAlaSprThrAlaGlnGlyA-rGValLeuGlyLysrTyrrVal 39
QY 154 ACTGTG-----CTGGGAAGCCCTGTGCTGTGAACGTGTTCTTGAGAGTCCCTTGTCT 207
Db 40 SerleuGlnGlyLeuAlaPheThrGlnPrrValAlaValAlaPheleuGlyValPrrPheAla 59
QY 208 GCTTCCCGCGCTGGGATCCCTGCGATTACGAACCGCGAGCTGCATCCCGCTGGAGTAA 267
Db 60 LysPrrProleuGlnGlySerleuA-rPheAlaPrrProGlnPrrAlaGluPrrTrpSerPhe 79

QY 268 TTGCAGAGCCACCTCTACCTTAATTGTGCTCCAGAACTCA-----GAGTGGCTG 321
Db 80 VallyrAasnThrTrhSerTyrrProPrometCysGlnAsePrrProValAlaGlnGlnMet 99
QY 322 CTCTTAGAT-----CAACACATGCTCAAGTGCATTACCCGAAA 360
Db 100 ThrSerAsePrrPheThrAasnPheThrGlyLysGlnA-rGLeuThrLeuGluPhe----- 117
QY 361 TTCCGAGGTGCAGAAAGTGTCTTACCTGAACATCTTAGGCGCCCGCCACCGCATACA 420
Db 118 -----SerGlnAsePrrCysleuTyrrLeuAasnleTyrrThrrProAlaAsePrrLeuThrLys 134
QY 421 GAGTCAAGCTCCCGCTTGTGATGCTTCCAGAGAGTGCCTTCAAGACTGCTGCTCAGCC 480
Db 135 ArgGlyA-rGLeuPrrProValMetValTrpLethrGlyGlyLeuValLeuGlyGlyAla 154
QY 481 TCCATCTTGAATGGGTCCGCGCTGCTGCTCATAGAGAC-----GTGCTGTTGTGCTG 534
Db 155 PrometTyrrAsePrrGlyValValleuAlaAlaThrGlnAasnPheThrValValValAla 174
QY 535 GTCCAGTACCGGCTTAGAAATATTGGTTTCTTACCAACATGGAGTACAGCATGCTCCGGGG 594
Db 175 LLeuGlnTyrrA-rGLeuGlyLethrGlyPhePheSerThrGlyAsePrrGlnHisSerA-rGly 194
QY 595 AACTGGGCTTCAAGAGACAGAGTGGCTGCTGTCTGCTGCTGCTCAAGAAACATCGAGTTC 654
Db 195 AasnTrpGlnHisLeuAsePrrGlnValAlaAlaLeuHisTrpValGlnGlnAasnHisAlaAasn 214
QY 655 TTGCGTGGGAGACCCGAGCTCTGTGACCATCTTGGCGAGTCC-----GCCGAGCCATA 708
Db 215 PheGlyGlyAsePrrProGlnSerValThrLethrPheGlyGlnSerPheThrAlaGlyGlyLys 234
QY 709 AGCTTTTCTACTTAACTCTCTCCATGGCCAAAGCTTATCCCAAAAGCATCTG 768
Db 235 SerValSerValleuValleuSerPrrProleuAlaLysAasnleuPrrPheHisA-rGAlaHis 254
QY 769 GAGAGTGGGCTGAGCATCATCTTACCTGAGAGGCCCATGATTAGAAAGTGAAGAGAC 828
Db 255 GlnSerGlyValAlaLeuThrValAlaLeuValA-rGlyAsePrrMetLysAlaAlaLys 274
QY 829 CTGACAGTGTGTCACATTCTGTGTGTAACAATGCTGACATCTGAGGCCCTGCTGAGG 888
Db 275 GlnHisAlaValleuAlaGlyCysleuThrThrThrrSerAlaValAlaPheThrPheValHis 294
QY 889 TGCCTGAGAGCAAAACCTCCAAAGAGCTGTGACCTCCAGCCAGAAACAAGCTTTC 948
Db 295 CysleuA-rGlnLysSerGlnAsePrrLeuLeuAsePrrLeuTyrrMetLysPheLeu 314
QY 949 ACT-----CGAGTGGTGTAT 963
Db 315 ThrleuAsePrrPheHisGlyAsePrrGlnA-rGlnSerHisPrrPheLeuPrrThrValValAsePrr 334
QY 964 GGTGCTTCTTCTTAAATGAGCTCTAGATCTATTGTCTCAGAAA-----GCATTTAAA 1017
Db 335 GlyValleuLeuPrrProLysMetPrrGlnGlnLeuAlaGlnLysAsePrrPheThrPheAasn 354
QY 1018 GCATTCCTTCCATCATCGAGTCAATTAACACAGAGTGTGCTTCCGTGCTGCT----- 1071
Db 355 ThrValPrrTyrrLethrValGlyLethrAasnLysGlnLysPheGlyTyrrPrrLeuPrrThrMet 374
QY 1072 -----ATGAAGAGAGCTCCCTGAGATCCAGT 1098
Db 375 MetGlyPhePrrProleuAsePrrGlnGlyLysleuAsePrrGlnLysThrAlaThrSerleuLeuTrp 394
QY 1099 GGTCCAAACAAGTCCCTTGCCTTCATCTGATACAAAACATCTGCAATCCGCTCAG 1158
Db 395 LysSer-----TyrrProLethrAasnle-----PrrGln 404
QY 1159 TATTTCACCTTGGCTGAATGAATACTTCATACACAGAC-----TCC 1203
Db 405 GlnleuThrPrrProValAlaThr-----PheThrAsePrrTyrrLeuGlyGlyThrAsePrr 422
QY 1204 CTGACTGAATCCGAGACAGTCTTCTGAGACTTGTGAGATGTGTCTTGTGTGCTCCT 1263


```

Db      155  PrometGlyValValLeuAlaAlaAlaSerLeuSerThrValMetLysPheTrpAla 174
Qy      535  GTCCAGTACCGGCTAGGAATATTGGTTTCTTCAACAATGGGATACAGATGCTCCGGG 594
Db      175  TleGlnTyArgLeuGlyLleTrpGlyPhePheSerThrGlyAspGlnHisSerArgGly 194
Qy      595  AACTGGGCTTCAAGGACGAGGGCTGTGCTGGGTCCAGGAAGAAACATCGAGTTC 654
Db      195  AantTrpGlnHisLeuSerPheGlnValAlaAlaLeuHisTrpValGlnMetLysAlaAla 214
Qy      655  TTCGGTGGGACCCGAGCTGTGACCATCTTTGGGAGTCC-----GCGGAGCCATA 708
Db      215  PheGlyGlyAspProGlySerValThrLlePheGlyGlnSerPheThrAlaGlyGln 234
Qy      709  AGTGTTTAGTCTTAATAGTCTGTCCATGGCCAAAGGCTTTATCCAAAGCCATCATG 768
Db      235  SerValSerValLeuLeuValLeuSerProLeuAlaLysAsnLeuPheHisArgAlaLieser 254
Qy      769  GAGAGTGGGGTGGCCATCCCTTAACCTGAGGCCCCATGATTAAGAAAGTGAGAC 828
Db      255  GlnSerGlyValAlaLeuThrValAlaLeuValArgLysAspMetLysAlaAlaLys 274
Qy      829  CTGCAAGTGGTTCGACATTTCTGGGTAAACAATGCTCAGACTCTGAGGCCCTGTGAG 888
Db      275  GlnLleAlaValLeuAlaGlyCysLysThrThrThrSerAlaValPheThrPheValHis 294
Qy      889  TGCCTGAGGACAAACCCCTCCAAAGAGCTGTGAGCCCTGACCCGAAACAAAGCTTTC 948
Db      295  CysLeuArgGlnLysSerGlnAspGlnLeuAspLeuThrLeuLysMetLysPheLeu 314
Qy      949  ACT-----CGAGTGGTGAAT 963
Db      315  ThrLeuAspPheHisGlyAspGlnArgGlnLysSerPheLeuProThrValAlaAsp 334
Qy      964  GGTCCTTTCTTCTTAATGAGCCCTAGATTAATTTCTCAGAAA-----GCATTTAA 1017
Db      335  GlyValLeuLeuProLysMetProGlnGlnLleAlaGlnLysAspPheThrPheAsn 354
Qy      1018  GCAATTCCTTCAATCATCGAGTCAATAACAAGAGTGGCTCTGCTGCTCT 1071
Db      355  ThrValProTyxLleValGlyLleAsnLysGlnGlnPheGlyTrpLeuLeuProThrMet 374
Qy      1072  -----ATGAGAGGCTCTGAGATCCCTCAT 1098
Db      375  MetGlyPheProLeuSerGlnGlyLysLeuAspGlnLysThrAlaHisSerLeuLeuTrp 394
Qy      1099  GGCTCCAAACAGTCCCTTCCCTCATCTGATACAAACATCCTGACATCCCGCTCAG 1158
Db      395  LysSer-----TyrProLleAlaAsnLle-----ProGln 404
Qy      1159  TATTGGACCTTGTGGTAATGATTAATCTTCATGACAGAC-----TCC 1203
Db      405  GlnLeuThrProValAlaAlaThr-----PheThrAspLysTyrLleGlnGlyLysAspAsp 422
Qy      1204  CTGACTGAATCCGAGACAGTCTTCTGACTCTTGGAGTGTCTTCTTGAGTCCCT 1263
Db      423  ProValLysLysLeuAspLeuPheLeuAspLeuMetGlyAspValAlaPheGlyValPro 442
Qy      1264  GCATGATACAGACATGATATACAGAGATGCTGTGACACTGTCTTATTAAGTTT 1323
Db      443  SerValThrValAlaArgGlnHisArgAspAlaGlyAlaProThrTyxMetTyxGlnPhe 462
Qy      1324  CGGACCCGGCTCAGTCTTTAAGAC-----ACGACCCGGCTTTTGTCAAGCCGAC 1377
Db      463  GlnTyArgProSerPheSerSerAspLysPheThrLysProLysPheThrAlaLleGlyAsp 482
Qy      1378  CAGCTGATGAAGTCCGCTTGTGTTGGTGGTGGCTTCTGAGGGGAGCAATGTGTAAG 1437
Db      483  HisGlyAspGlnLlePheSerValPheGlyPheProLeuLeuLysGlyAsp----- 499
Qy      1438  TTCGAAGGCCACCGAGAGGAGAAATTACTGAGCCGAGAAAGTGAATTAATCTGGGCT 1497

```

```

Db      500  -----AlaProGlnGlnGlnValSerLeuSerTyxThrValMetLysPheTrpAla 516
Qy      1498  ACCTTGTCTGCAACCGGGAATCTTAATGGGACGACCTGTCTGTGGCCA-----GCT 1551
Db      517  AsnPheAlaArgSerGlyAsnProAsnGlnGlnGlyLysPheProHisTrpProPheThrMet 536
Qy      1552  TATAATCTGATCGACAGTACCTCCAGCTTGGAATGATGAGCTCTCGAGACAGACTTC 1611
Db      537  TyxAspGlnGlnGlnGlyTyxLleGlnLleGlyValAsnThrGlnAlaAlaLysArgLeu 556
Qy      1612  AAAGAACCGGGTGGATTTTGGACCGACCAATC 1647
Db      557  LysGlyGlnGlnValAlaPheTrpAsnAspLeuLeu 568

RESULT 12
US-09-264-737-1
; Sequence 1, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruf, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; FILE REFERENCE: Expression of Bacterial Enzymes
; CURRENT FILING DATE: 1999-03-09
; EARLIER FILING DATE: 1998-03-10
; EARLIER APPLICATION NUMBER: 60/077,377
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Rabbit
US-09-264-737-1

Alignment Scores:
Pred. No.: 7,99e-93 Length: 539
Score: 1034.50 Matches: 234
Percent Similarity: 55.05% Conservative: 71
Best Local Similarity: 42.24% Mismatches: 176
Query Match: 32.63% Indels: 73
DB: Gaps: 13

US-10-023-515-3 (1-1746) x US-09-264-737-1 (1-539)
Qy      94  CCTTGTCTGGAAGGCCACAGAGAAACACAGCTGGAGTGAATTCAGGCAAGCAAGTC 153
Db      2  ProSerAlaPro---ProValAlaAspThrValLysGlyLysValLeuGlnLysPheVal 20
Qy      154  ACTGTGCTGGGAAGCCCTGTGCTGTGAACGTCTTCTGGAGTCCCTTGTGCTGCC 213
Db      21  SerLeuGlnGlyPheAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 40
Qy      214  CCGTGGAGTCCCTGGATTTACGAACCCGAGCCGTGATCCCGCTGGATTAATCTGCA 273
Db      41  ProLeuGlySerLeuAspPheAlaProProGlnProAlaGlnLysSerTrpSerHisValLys 60
Qy      274  GAAGCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTTAGATCAA 333
Db      61  AsnThrThrSerTyxProPheMetCysSerSerAspAla-----ValSerGly 76
Qy      334  CACATGCTC-----AAGGTCATTAACCC---AAATTCGAGAGT 369
Db      77  HisMetLeuSerGlnLeuPheThrAsnArgLysGlnAsnLleProLeuLysPhe----- 94
Qy      370  TCAGAAAGTGCCTCTAATCTGAACATCTAATGGCTGCGCCAGCCGATACAGGCTCCAG 429
Db      95  SerGlnAspCysLeuTyxLeuAsnLleTyxThrProAlaAspLeuThrTyxArgGlyAsp 114
Qy      430  CTCCTCGCTTGTGTGTGCTCCAGAGAGTGCCTTCAGACTGGCTCAGCTTCATCTTT 489
Db      115  LeuProValMetValTrpLleHisGlyGlyGlyLeuMetValGlyGlyAlaSerThrTyx 134

```



```

QY 433 CCGCTCTGTGTGCTCCAGAGAGTCCCTCAAGACTGCTCAGCTCCATCTTGTAT 492
DB 21 ProValmetValTPrIleHisGlyGlyLeuMetValGlyAlaLaserThrTyraP 40
QY 493 GAGTCCGCTGCTGCTGCTATGAGAGCTGCTGTGTGTGCTGCTCAAGTACCGGCTAGA 552
DB 41 GlyLeuAlaLeuAlaHisGlyLeuMetValValThrIleGlyTrzArgLeuGly 60
QY 553 ATATTGTGTTCTTCCACAGAGATGAGATGCTCCGGGGAGCATGGGCTTCAAGAGC 612
DB 61 IletPrGlyPhePheSerThrGlyAspGlyHisSerArgGlyAsnTrpGlyHisLeuAsp 80
QY 613 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
DB 81 GluValAlaAlaLeuAlaGlyTrpValGlnAspAsnIleAlaSerPheGlyGlyAsnProGly 100
QY 673 TCTGTGACCATCTTTGGCGAGTCCCGGAGGCCATTAAGTGTCTTCACTTACTGTCT 732
DB 101 SerValThrIlePheGlyGlySerAlaGlyGlyGlySerValSerValLeuValLeuSer 120
QY 733 CCGATGCGCCAAAGGCTTATTCACAAAGCATCAGAGAGAGTGGGCTGCTCATCCCT 792
DB 121 ProLeuAlaLysAsnLeuPheHisArgAlaIleSerGlySerGlyValAlaLeuThrSer 140
QY 793 TACCTGAGAGGCCATGATTATGAGAAG-----AGTGAAGACCTGCAAGTGGTGCACAT 846
DB 141 ValLeuValLysGlyAspValLysProLeuAlaGlyGlnIleAlaIleThrAla--- 159
QY 847 TTCTGTGTAAACATGACCTGAGACTGAGGCTGAGGCTGAGGCTGAGGACAAACCC 906
DB 160 -----GlyCysLysThrThrThrSerAlaAlaMetValHisCysLeuAspGlyLysThr 177
QY 907 TCAGAGAGCTGCTGACCTCAGCCAGAA----- 936
DB 178 GluGluGluLeuLeuGluThrThrLeuLysIleGlyAsnSerTyLeuTrpThrTyraG 197
QY 997 -----ACAAAGCTTTCACTCGAGGTGATGAGGTGCTTCTTCTTCTTAT 981
DB 198 GluThrGlnAspGlySerThrLeuLeuGlyThrValIleAspGlyMetLeuLeuLys 217
QY 982 GAGCTCTAGATCTATTGCTGAGAAAGCATTTAAAGCAATTCCTTCATCAGGAGTC 1041
DB 218 ThrProGluGluLeuGlnArgGlnArgAsnPheHisThrValProTyMetValGlyLe 237
QY 1042 AATAACACAGAGTGGCTTCTGCTGCTGCTATGAGAGAGGCTCCGAGATCTCAGTGGC 1101
DB 238 AsnLysGlnGluPheGlyTrpLeuIlePrometGlnLeuMetSerTyzProLeuSerGlu 257
QY 1102 TCC-----AACAAAGCTTCCCTCCATCTGATTAACAAC-----ATCCTGCAC 1146
DB 258 GlyGlnLeuAspGlnLysThrAlaMetSerLeuLeuGlySerProIleProLeuPheAla 277
QY 1147 ATCCGCTCAGTATTATTGACCTTGGGCTTAATGATTAATCTTCATCAGACACTCCCTG 1206
DB 278 IleAlaLysGluLeuIleProGluAlaThrGluLysTyLeuGlyGlyThrAspAspThr 297
QY 1207 ACTGAATCCGAGACAGCTTCTGAGCTTCTGAGATGTGTTCTTGTGTGCTCCTGCA 1266
DB 298 ValLysLysLeuAspLeuIleLeuAspLeuIleAlaAspValMetPheGlyValProSer 317
QY 1267 CTGATACACAGCTGATATACAGAGATGCTGTGTCACCTGTACTCTTCAATGAGTTTGG 1326
DB 318 ValIleValAlaArgAsnHisArgAspAlaGlyAlaProThrTyMetTyArgLysPheGln 337
QY 1327 CACCGGCTCAGTGTGTTGAAGACACGAGAGCGGCTTTGTCAAGCCGACACGCTGAT 1386
DB 338 TyraArgProSerPheSerSerAspMetLysProLysThrValIleGlyAspHisGlyAsp 357
QY 1387 GAAGTCCCGCTTGTGTGCTGCTGCTTCTGTAAGGGGAGCATTTGTATGTTCAAGGA 1446
DB 358 GluLeuPheSerValPheGlyAlaProPheLeuLys-----GluGly 371
QY 1447 GCCACGAGAGAGAGAAGTTACTGAGCCGAGAGATGATGAATATCTGGGCTACTTGTGCT 1506

```

```

DB 372 AlasSerGluGluGluGlnIleArgLeuSerLysMetValMetLysPheTrpAlaAsnPheAla 391
QY 1507 CGAACCGGGAATTCCTAATGGGAGACACCTGCTCTGTGCGCCAGTATTATCTGACTGAG 1566
DB 392 ArgAsnGlyAsnProAsnGlyLysGlyLeuProHisTrpProGluTyraHisGlnLysGlu 411
QY 1567 CAGTACTCTCAGCTGACTTGAACATGAGCCTCGACAGACAGACTCAAGAACCGCGGCTG 1626
DB 412 GlyTyrlLeuGlnIleGlyAlaAsnThrGlnAlaIleGlnLysLeuLysAspLysGluVal 431
QY 1627 GATTTTGGACGACG 1641
DB 432 AlaPheTrpTrpAsn 436

RESULT 14
US-08-446-100-28
/ Sequence 28, Application US/08446100
/ Patent No. 6001625
/ GENERAL INFORMATION:
/ APPLICANT: Broomfield, Clarence A
/ APPLICANT: Willard, Charles B
/ APPLICANT: Lockridge, Oksana
/ TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Hendricks and Assoc.
/ STREET: 9669 A Main Street, P.O. Box 2509
/ CITY: Fairfax
/ STATE: VA
/ COUNTRY: US
/ ZIP: 22031
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/446,100
/ FILING DATE: 19-MAY-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hendricks, Glenna
/ REGISTRATION NUMBER: 32,535
/ REFERENCE/DOCKET NUMBER: broomfield
/ TELEPHONE: (703) 425-4250
/ TELEFAX: (703) 425-2767
/ INFORMATION FOR SEQ ID NO: 28:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 454 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ANTI-SENSE: YES
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: human esterases
/ US-08-446-100-28

Alignment Scores:
Pred. No.: 8.37e-84 Length: 454
Score: 942.50 Matches: 200
Percent Similarity: 59.55% Conservative: 65
Best Local Similarity: 44.94% Mismatches: 149
Query Match: 29.73% Indels: 31
DB: 3 Gaps: 6

US-10-023-515-3 (1-1746) x US-08-446-100-28 (1-454)
QY 373 GAAGACTGCTCTACTGAAATCTATGCGCTGCGCCAGCGAGTACAGGCTTCAAGCTC 432

```


US-10-023-515-3 (1-1746) x US-08-446-100-30 (1-454)

[illegible]

QY	1387	GAAC	TCCGCTTTGGTTCGGGCGGCGCTTCGTAAGGGGGGACATTTGATTGTCGAAGGA	1446
		:::		
Db	358	GIUueheserValPheGlyAlaProPheuuys-----	GIUGly	371
QY	1447	GCCACGGAGGAGAGAAATTACTGAGCCGGAGAGATGATGAATACTGGCTACTTTGCT		1506
		:::	:::	
Db	372	AlaserGIuGIuGIuIleArgLeuSerZlyMetValMetLysPheTrpAlaAsnPha		391
QY	1507	CGAACCGGGAAATCTTATGGAACGACCTGTCTGTGGCAGCTTATATCTGACTGAG		1566
Db	392	ArgAsnGIyAsnProAsnGIyLysGlyLeuProHisTrpProGIuTrpAsnGlnLysGlu		411
QY	1567	CAGTACCTCCAGCTGGAATTGAACATGAGGCTCGGACAGAGACTCAAGAAGCCGGGGTG		1626
Db	412	GIyTrpLeuGlnIleGlyAlaAsnThrGlnAlaAlaGlnLysLeuSAsPlyGluVal		431
		:::	:::	
QY	1627	GATTTTGGACCAAC	1641	
Db	432	AlaPheTrpIuTrAsn	436	

Search completed: June 15, 2005, 08:57:41
Job time : 53.3868 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 08:53:20 ; Search time 120.306 Seconds

(without alignments)
11126.674 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 3170
Sequence: 1 atgcacacaggaacttacttc.....ttctcttgcctctctga 1746

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 3420798

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=x1h
-O=/cgn2_1/USPTO_epool/6664091/runat_14062005_133501_9433/app_query.fasta_1.4238
-DB=Published Applications AA -QMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USRR=6664091.@CEN_1.1.295.@runat_14062005_133501_9433
-ICPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US16_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	97.1	581	13	US-10-023-515-2
2	3079	97.1	581	15	US-10-674-636-2
3	3079	97.1	581	16	US-10-757-262-46
4	2915	92.0	575	15	US-10-451-168-91
5	2900	91.5	642	15	US-10-433-256-10
6	2607.5	82.3	581	15	US-10-451-168-92
7	2606	82.2	525	15	US-10-094-749-2275
8	2597.5	81.7	581	15	US-10-114-270-196
9	2588.5	81.9	618	15	US-10-381-898-2
10	2454	77.4	469	15	US-10-104-074-2219
11	1985	62.6	542	16	US-10-233-933A-2
12	1985	62.6	542	16	US-10-233-933A-4
13	1780	56.2	356	15	US-10-451-168-93
14	1230	38.8	549	9	US-09-925-298-689
15	1230	38.8	549	14	US-10-102-806-689
16	1230	38.8	550	15	US-10-267-756-5
17	1230	38.8	559	16	US-10-858-271-28
18	1186.5	37.4	554	9	US-09-895-860-4
19	1186.5	37.4	554	15	US-10-377-072-4
20	1186.5	37.4	554	16	US-10-377-072-4
21	1179.5	37.2	571	10	US-09-931-836-23
22	1179.5	37.2	571	11	US-09-833-245-1090
23	1179.5	37.2	571	13	US-10-036-342-23
24	1179.5	37.2	571	13	US-10-036-041-23
25	1179.5	37.2	571	14	US-10-028-072-542
26	1179.5	37.2	571	14	US-10-035-855-23
27	1179.5	37.2	571	14	US-10-140-808-542
28	1179.5	37.2	571	14	US-10-121-049-542
29	1179.5	37.2	571	14	US-10-123-904-542
30	1179.5	37.2	571	14	US-10-140-470-542
31	1179.5	37.2	571	14	US-10-140-746-542
32	1179.5	37.2	571	14	US-10-176-921-542
33	1179.5	37.2	571	14	US-10-176-921-542
34	1179.5	37.2	571	14	US-10-227-884-210
35	1179.5	37.2	571	14	US-10-036-214-23
36	1179.5	37.2	571	14	US-10-137-865-542
37	1179.5	37.2	571	14	US-10-140-474-542
38	1179.5	37.2	571	14	US-10-035-719-23
39	1179.5	37.2	571	14	US-10-142-431-542
40	1179.5	37.2	571	14	US-10-143-114-542
41	1179.5	37.2	571	14	US-10-230-163-210
42	1179.5	37.2	571	14	US-10-036-160-23
43	1179.5	37.2	571	14	US-10-230-338-210
44	1179.5	37.2	571	14	US-10-142-419-542
45	1179.5	37.2	571	14	US-10-218-631-210

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: S10s-Santiago, Immaculada
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279, 508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581

TYPE: PR1
ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:

Pred. No.:	4,94e-267	Length:	581
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	97.13%	Indels:	0
	13	Gaps:	0

US-10-023-515-3 (1-1746) x US-10-023-515-2 (1-581)

```
QY 1 ATGCCACGGGAACTTACTGCTGCTTGCATGAGTGGCTTTTCCATGATCCAGCCC 60
DB 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPheHeuLeuLeuGlnPro 20
QY 61 CTGTTGGGACACAGACAGTGGGAAAAAAGCTGGGCTTCTGCTGAAGGCGCAGAGGAAC 120
DB 21 LeuLeuGlyHisArgGlnTrpGlySerThrGlyProSerAlaGlnGlyProGlnArgAsn 40
QY 121 ACCAGGCTGGGATGGATTGAGGAGCAAGTCACTGTGCTGGGAAAGCCCTGCTGCTG 180
DB 41 ThrArgLeuGlyTrpIleGlnGlySerGlnValThrValLeuGlySerProValProVal 60
QY 181 AACGTTCTCTGGAGTCCCTTTGCTGCTCCCGGCTGGGATCCCTGAGATTATCGAAC 240
DB 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheTrpAsn 80
QY 241 CCGAGCCCTGATCGCCCTGGGATTAATTGCGAGAGCCACCTCCTACCTTAATTTGTGC 300
DB 81 ProGlnProAlaSerProTrpAsnLeuArgGlnAlaThrSerTrpProAsnLeuGly 100
QY 301 CTCGGAAGTCAAGTGGTCTCTTAAATCAACACATGCTCAAGTGATTAACCGGAAA 360
DB 101 LeuGlnAsnSerGlnTrpLeuLeuAspGlnHisMetLeuIleValHisTrpProGly 120
QY 361 TTGGAGTGTCAAGAGACTGCTTACCTGAAATCATCTATGCGGCTGGCCGCGATACA 420
DB 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
QY 421 GGGCTCAAGCTCCCGCTCTTGATGTGTTCCAGAGAGTGCCTTCAAGACTGAGCC 480
DB 141 GlySerTrpLeuProValLeuValTrpPheProGlyAlaPheLeuSerThrGlySerAla 160
QY 481 TCCATCTTGAATGGTCCGCTGCTGCTTATGAGAGCGTGTGTTGCTGCTCCAG 540
DB 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValAlaGln 180
QY 541 TACGGGCTAGGAATATTTGGTTCTTCCACACATGGGATCAGATGCTCCGGGAACTGG 600
DB 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHisAlaProGlyAsnTrp 200
QY 601 GCGTTCAAGAGACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 AlaPheIleAspGlnValAlaAlaLeuSerTrpValGlnIleAsnIleGlnPheGly 220
QY 661 GGGGACCCCAAGTCTGTGACCATCTTTGGCGAGTCCGCGGAGCCATAAGTCTTCTACT 720
DB 221 GlyAspProSerSerValThrIlePheGlyIleSerAlaGlyAlaIleSerValSerSer 240
QY 721 CTTATACTGTCTCCCATGGCCAAAGGCTTATCCCAAAAGCATGAGTGAAGTGGGGG 780
DB 241 LeuIleLeuSerProMetAlaIleAspGlyLeuPheHisIleValAlaIleMetGlnSerAla 260
QY 781 GCCATCATCCCTTACCTGAGGCGCCATGATTAAGAAGAGTGAAGAGCACTGAGGTGTT 840
DB 261 AlaIleIleProGlyLeuGlnIleHisAspTrpGlnIleSerGlnAspLeuGlnValVal 280
QY 841 GCACATTTCTGTGGTGAACATGCGTCAAGACTGAGGCGCTGCTGAGGTGCTGAGGACA 900
DB 281 AlaHisPheCysGlyAlaAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
```

```
QY 901 AAACCTCCAGGAGCTGTGACCTTCAGCCGAGAAAACAAGTCTTTCACCTGAGTGGTT 960
DB 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysTrpLysSerPheThrArgValAla 320
QY 961 GATGGTCTTCTTCTTAATGAGCCCTAGATCTATTGCTCAGAAAACATTTAAAGCA 1020
DB 321 AspGlyAlaPhePheProAsnGlnTrpLeuAspLeuLeuSerGlnIleValAlaPheIleValAla 340
QY 1021 ATTCCTTCATCATGTGAGTCAATTAACCAAGAGTGGCTTCTGCTGCTTATGAAGAG 1080
DB 341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLysGln 360
QY 1081 GCTCCTGAGATTCCTGAGTGGCTCCAAAGTCCCTTGCCTCCATCTGATACAACATC 1140
DB 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIle 380
QY 1141 CTGACATCCCGCTCAGTATTTGGACCTTGCTGCTAAATGAATCTTCATGACAAAGAC 1200
DB 381 LeuHisIleProProGlnTrpLeuHisIleuValAlaAsnGlnTrpPheHisAspLysHis 400
QY 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGACTTGTGGAGATGTGTTCTTGTGGTC 1260
DB 401 SerLeuThrGlnIleArgAspSerLeuAspLeuAspLeuGlyAspValPhePheValAla 420
QY 1261 CCTGACTGATCAACAGTCTGATATCAAGAGATGCTGGTGCACCTGCTACTTCTATGAG 1320
DB 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGln 440
QY 1321 TTTGGGACCGGCTCAGTGTCTTTGAAGACAGAAAGCGGCTTTGTCAAGCCGACAC 1380
DB 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValIleAspHis 460
QY 1381 GCTGATGAAGTCCGCTTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 461 AlaAspGlnValAlaArgPheValPheGlyAlaPheLeuLysGlyAspIleValMetPhe 480
QY 1441 GAAGAGCCACGAGAGAGAGAAATTAATGAGCCGAGAGATGAATTAATGAGGCTACC 1500
DB 481 GlnGlyAlaThrGlnGlnGlnGlnIleLeuLeuSerArgLysMetMetLysTrpAlaThr 500
QY 1501 TTTGCTGAACCGGAGATCTTAATGGAACGACTGTCTGTGGCCAGCTTAATATGTG 1560
DB 501 PheAlaArgTrpGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
QY 1561 ACTGAGCATACCTCCACTGAGACTTGAACATGAGCCCTCGGACAGACTCAAGAAACG 1620
DB 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
QY 1621 CGGATGATTTTGGACACGACCATCCCTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
QY 1681 AGTCTCTTTTCTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
DB 581 Pro 581
```

RESULT 2
US-10-674-636-2
; Sequence 2, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674,636
; CURRENT FILING DATE: 2003-09-29

; PRIOR APPLICATION NUMBER: US/10/023, 515
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 60/256, 369
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: 60/279, 508
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-674-636-2

Alignment Scores:

Pred. No.:	4,94e-267	Length:	581
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	97.13%	Indels:	0
DB:	15	Gaps:	0

US-10-023-515-3 (1-1746) x US-10-674-636-2 (1-581)

```

QY 1 ATGCCACAGGACTTACTTCATCTGCTTCAACAATGGCTTTTCTGATTTCTCAGCC 60
DB 1 MetProGlnGlyLeuTherSerSerAlaSerGlnTrpCyPhePheLeuIleLeuGlnPro 20
QY 61 CTGTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGACACAGAGAAC 120
DB 21 LeuLeuGlnYH1aArgGlnTrpGlyLeuThnGlyProSerAlaGlnGlyProGlnArgAn 40
QY 121 ACCAGGCTGGGATTTGATTTCAAGGACAGTCACTGTCTGTGGAAAGCCCTGCTGTG 180
DB 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
QY 181 AACGTGTCTCGAGATCCCTTTGTGCTGCCCGCTGGAGATCCCTGGCATTTACGAAC 240
DB 61 AsnValPheLeuGlyValProPheAlaAlaProPheLeuGlySerLeuArgPheThrAn 80
QY 241 CCGAGGCTGATCGCCCTGGGATTAATTGCGAAGAACCACTCTTAATTTTGTGC 300
DB 81 ProGlnProAlaSerProTrpPhePheLeuArgGlnAlaThrSerTrpProLeuLeuGly 100
QY 301 CTCGCAAGACTCAGAGTGGCTGCTTCAATCAACACATGCTCAAGGTGATTAACCGGAA 360
DB 101 LeuGlnAsnSerGlyTrpLeuLeuAspGlnH1aMetLeuLysValH1aTrpProLys 120
QY 361 TTGGAGTGTGAGAGACTGCTTACCTGAAATCTATGCGCTGCGCAGCCGATACA 420
DB 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnH1aTrpAlaProAlaH1aAlaAspTr 140
QY 421 GGCTCCAAAGCTCCGCTTGTGTGTGTCCAGAGAGTGCCTTCAAGACTGCTCAGCC 480
DB 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
QY 481 TGCATCTTGAATGGGTCCGCTGCTGCTTATGAGAGAGTGTGTGTGTGCTGCTCAG 540
DB 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValAlaValGln 180
QY 541 TACCGGATGAGAAATTTGTTCTTCAACATGGGATCAGATGCTCCGGGAACTGG 600
DB 181 TyrArgLeuGlyIlePheGlyPhePheTrpTrpPheGlnH1aAlaProGlyAsnTrp 200
QY 601 GCCTTCAAGACCAAGTGTGCTGTCTGTCTGAGGCTCAGAGAAACATCGATTTCTCGGT 660
DB 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnH1aGlnPheGly 220
QY 661 GGGGACCCCAAGCTCTGTGACATCTTTGGCAGAGTCCGCGGAGCCATAAGTTTCTACT 720
DB 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
QY 721 CTTATATCTGTCTCCATGGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGTG 780
  
```

```

DB 241 LeuIleLeuSerProMetAlaLysGlyLeuPheH1aLysAlaIleMetGlnSerGlyVal 260
QY 781 GCCATCATCCCTTACTCGTAGGCCCCATGATTTAGAAAGAGAGAGACCTGCAAGTGGTT 840
DB 261 AlaIleIleProTyLeuGlnAlaH1aAspTrpGlnLysSerGlnAspLeuGlnVal 280
QY 841 GCACATTTCTGTGTGAACAATGCGTCAAGTCAAGTGGCTGCTGAGTGTGAGGAGCA 900
DB 281 AlaH1aPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuAlaGln 300
QY 901 AAACCTCCAAAGAGCTGTGACCTTCAAGCCAGAAACAAAGCTTTTCACTGAGTGGTT 960
DB 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
QY 961 GATGGTCTTTCTTTCTTAATGAGCTCTAGATCTATTTGTCTCAAGAAACATTAAAGCA 1020
DB 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
QY 1021 ATTCCCTTCATGATCGAGTCAATAACACAGAGTGGCTTCCCTGCTGCTATGAAGAG 1080
DB 341 IleProSerIleIleGlyValAsnAsnH1aGlnCysGlyPheLeuLeuProMetLysGln 360
QY 1081 GCTCTGAGATCTCAGTGGCTCAACAAAGTCCCTTGCCTTCATGTATACAAACATC 1140
DB 361 AlaProGlnIleLeuSerGlySerLeuLysSerLeuAlaLeuH1aLeuIleGlnAsnIle 380
QY 1141 CTCGACATCCCGCTCAGATTTTGAACCTTTGGGCTTAAGATCTTCCATGACAGACAC 1200
DB 381 LeuH1aIleProGlnTrpLeuH1aLeuValAlaAsnGlyTrpPheH1aAspLysH1a 400
QY 1201 TCCCTGATGAATATCCGAGACAGTCTTGTGACTTGTGAGATGTGTTCTTGTGGTGC 1260
DB 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 420
QY 1261 CCTGCACTGATCAACAGCTGATATCAACAGATGCTGTGACCTTGTACTTATGAG 1320
DB 421 ProAlaLeuIleThrAlaArgTrpH1aAspArgAlaGlyAlaProValTrpPheTrpGln 440
QY 1321 TTTCGACACCGGCTCAGTGCCTTTGAAGACAGAGACCGGCTTTTGTCAAGCCGACAC 1380
DB 441 PheArgH1aArgProGlnCysPheGlnAspTrpTrpLysProAlaPheValLysAlaAspH1a 460
QY 1381 GCTGATGAATCGGCTTGTGTGTGTGCTGCTTGTGAGGAGGAGCATGTTATGTTTC 1440
DB 461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
QY 1441 GAAGAGCCACGAGAGAGAGAACTTACAGACCGGAGAGATGATAATCTGGCTTACC 1500
DB 481 GlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpTrpAlaThr 500
QY 1501 TTGTGCTGAACCGGGAATCTTAATGGAAACAGCTGTCTGTGTGCTGCTTAAATCTG 1560
DB 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAlaLeu 520
QY 1561 ACTGAGAGTACTCTCAGCTGAGTGAACATGAGGCTGTGACAGAGATCAAGAACCG 1620
DB 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
QY 1621 CCGGTGATTTTGGACACAGACATCCCTGATCTGTGCTGCTGCTGCTGCTGCTGCTCAG 1680
DB 541 ArgValAspPheTrpPheSerThrIleProLeuIleLeuSerAlaSerAspMetLeuH1a 560
QY 1681 AGTCTCTTTTCTTCTTAATCTTCTCTCTCTCTCTGCTCAGCCTTTTCTTTTGTGCT 1740
DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
DB 581 Pro 581
  
```

RESULT 3
 US-10-757-262-46

```
; Sequence 46, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2518, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MP103-00721RNMNMIM
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-46

Alignment Scores:
Pred. No.: 4,94e-267 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.13% Indels: 0
DB: 16 Gaps: 0

US-10-023-515-3 (1-1746) x US-10-757-262-46 (1-581)
QY 1 ATGCCACAGGAGCTTACTTCACTGCTTCAACATGAGTGTCTTTTCTGATTTCTCAAGCC 60
Db 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCySPhePheLeuLeuGlnPro 20
QY 61 CTGTTGGGACACAGCAGTGGGGAAGAACTGGGCTTGTGCTGAGAGGGGACAGAGAAAC 120
Db 21 LeuLeuGlnGlyAlaSerGlnTrpGlyPheSerAlaGlnGlyProGlnArgAsn 40
QY 121 ACCAGGCTGAGATGATTCAGGGCAAGCAAGTCACTGTGCTGAGAAAGCCCTGCTGTG 180
Db 41 ThrArgLeuGlyTrpLysGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
QY 181 AAGGTGTTCTGAGAGTCCCTTTGCTGCTCCCGGCTGGAGTCCCTGCGATTACGAAC 240
Db 61 AsnValPheLeuGlyValPheProPheAlaAlaProPheLeuGlySerLeuArgPheThrAsn 80
```

```
QY 241 CCGACGCTGATCGCCCTGGAGATTAATTGACGAGAGACCACTCTACCTAATTGTC 300
Db 81 ProGlnProAlaSerProTrpAspAsnLeuArgValAlaThrSerTrpProAsnLeuCyS 100
QY 301 CTCAGAACTCAGAGTGTGCTCTTGAATCAACACATGCTCAGAGTGATTAATCCGAA 360
Db 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuValHisIleTrpLeuCyS 120
QY 361 TTCGAGAGTCAAGAGTCCCTTCACTGAAACATCTATGCGCCGCCACGCGCATCA 420
Db 121 PheGlyValSerGlnAspCySLeuValLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
QY 421 GGCCTCAAGCTCCCGCTGTTGCTGTGCTCCAGAGAGTGCTTCAAGACTGCTCAGCC 480
Db 141 GlySerTrpLeuProValLeuValTrpPheProGlyGlyAlaPheValThrGlySerAla 160
QY 481 TCCATCTTTGATGGTCCGCCCTGGCTGCTATAGAGACGTGCTGTTGTGCTCCAG 540
Db 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValValGln 180
QY 541 TACCGGCTAGAAATTTTGTTGTTCTTCAACACATGGATTCAGATCCCGGGAACTGG 600
Db 181 TyrArgLeuGlyLysPheGlyPhePheThrTrpTrpAspGlnHisAlaProGlyAsnTrp 200
QY 601 GCCTTCAAGACACAGTGGCTGCTGTCTGCTGGGTCCAGAAACATGAGATTTCCGCT 660
Db 201 AlaPheValAspGlnValAlaAlaLeuSerTrpValGlnValAsnIleGlnPheGly 220
QY 661 GGGGACCCGAGCTGTGACCAATCTTTGACGAGTCCGCGGAGCCATAGTGTTCAGT 720
Db 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
QY 721 CTTATACGTCTCCCATGCGCAAGGCTTATTCACAAAGCATATGAGAGTGGGGTG 780
Db 241 LeuIleLeuSerProMetAlaValSerGlyLeuPheHisIleValIleMetGlnSerIleVal 260
QY 781 GCCATCATCCCTTACCTGAGGCGCATATTAATGAGAGAGTGAAGACCTGACGCTGTT 840
Db 261 AlaIleIleProTrpLeuGlnAlaHisIleAspTrpGlnValSerGlnAspLeuGlnVal 280
QY 841 GCACATTTCTGTGTAAACATATGCTGACACTCTGAGGCTTGTGAGTGTCTGAGACA 900
Db 281 AlaHisPheCySValAsnAsnAlaSerAspSerGlnAlaLeuValArgCySLeuArgThr 300
QY 901 AAACCTTCAAGAGCTGCTGACCTTCAAGCAAGAAAGATCTTCACTCGAGTGGTT 960
Db 301 LysProSerTrpGlnLeuLeuThrLeuSerGlnValTrpLysSerPheThrArgVal 320
QY 961 GATGAGTCTTCTTCTTAATGAGGCTTGAATCTTATGCTCAGAAAGCATTTAAAGCA 1020
Db 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnValAlaPheValAla 340
QY 1021 ATTCTTTCATCATCAGATCAATTAACACAGAGTGTGCTTCTGTGCTTATGAAGAG 1080
Db 341 IleProSerIleIleGlyValAlaAsnAsnHisGlnCySValPheLeuLeuProMetCySVal 360
QY 1081 GCTCTGAGATTCCTCAGTGGCTCCCAAGATCCCTTGGCCCTCCATCTGATCAAAACATC 1140
Db 361 AlaProGlnIleLeuSerGlySerAsnLysSerIleValAlaLeuHisIleGlnAsnIle 380
QY 1141 CTGCACATCCCGGCTCAGTATTTTGACCTTGTGCTAATGAATCTTCCATGACAGAC 1200
Db 381 LeuHisIleProProGlnTrpLeuHisIleValAlaAsnGlnTrpPheHisIleAspValHis 400
QY 1201 TCCCTGACTGAATTCGAGACAGTCTTGTGAATCTTGTGAGATGTGTTCTTTGTGTC 1260
Db 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
QY 1261 CCTGACATGATCAACAGCTCGATATACAGAGATGCTGTGACCTGTCTTACTTCAATAG 1320
Db 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGln 440
QY 1321 TTTGGGACCGGCGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTTGTCAAGCCGACAC 1380
```

```

|||||
Db PhearghlaargProglincyphhegluabPthrllysProAlaPheValIysAlaaphis 460
QY 1381 GGTATGTAAGTCCGCTTTGTGTTCCGTGTGCTTCCTGAAGGGGACATTTGTAATGCTC 1440
Db 461 ALaabPgluValaIarPheValaPhegllyAlaPheLeuIysglYAspIleValaMetPhe 480
QY 1441 GAAAGAGCAACGAGAGAGAGAAATTAAGTACAGCCGGAATGATGAATAATAGGCTACC 1500
Db 481 GUGlYAlaThrlGUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGl 500
QY 1501 TTGTGTCGAACCGGGAATCTTAATGGAAAGAACCTGTCTGTGTGCGCAGCTTAATAATCTG 1560
Db 501 PheAlaargThrlGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpLeu 520
QY 1561 ACTGAGCAGTACCTCCAGCTGAGCTTGAACATGAGCTTGACAGACATCTCAAGAACCG 1620
Db 521 ThrUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGl 540
QY 1621 CGGTCGATTTTGGACGACGACATCCCTGATCTGTGCTGCTGCGACATGCTCCAC 1680
Db 541 ArgValaPheTrpThrlSerThrlleProleuIleleuSerAlaSerAspMetLeuHis 560
QY 1681 AGTCCTCTTCTCTTCACTTTCCTCTCTCTCTCTGACCTTCTTCTTCTTCTTGTGCT 1740
Db 561 SerProleuSerSerLeuThrPheleuSerLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
Db 581 Pro 581

```

RESULT 4

US-10-451-168-91

; Sequence 91, Application US/10451168

; Publication No. US20040091969A1

; GENERAL INFORMATION:

; APPLICANT: SMITHKLINE BEECHAM CORPORATION

; APPLICANT: SMITHKLINE BEECHAM P.L.C.

; APPLICANT: GLAXO GROUP LIMITED

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GPs0039

; CURRENT APPLICATION NUMBER: US/10/451,168

; PRIOR FILING DATE: 2003-11-12

; PRIOR APPLICATION NUMBER: PCT/US01/49232

; PRIOR FILING DATE: 2000-12-17

; PRIOR APPLICATION NUMBER: 60/256,710

; PRIOR FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: 60/257,048

; PRIOR FILING DATE: 2000-12-20

; PRIOR APPLICATION NUMBER: 60/260,482

; PRIOR FILING DATE: 2001-01-09

; PRIOR APPLICATION NUMBER: 60/264,922

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 60/266,797

; PRIOR FILING DATE: 2001-02-06

; PRIOR APPLICATION NUMBER: 60/276,988

; PRIOR FILING DATE: 2001-03-19

; PRIOR APPLICATION NUMBER: 60/281,535

; PRIOR FILING DATE: 2001-04-04

; PRIOR APPLICATION NUMBER: 60/289,622

; PRIOR FILING DATE: 2002-06-28

; NUMBER OF SEQ ID NOS: 110

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 91

; LENGTH: 575

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-451-168-91

Alignment Scores:

Pred. No.: 2,61e-252

Score: 2915.00

Percent Similarity: 97.54%

Length: 575

Matches: 553

Conservative: 3

```

Beet Local Similarity: 97.02% Mismatches: 6
Query Match: 91.96% Indels: 8
DB: 15 Gaps: 1
US-10-023-515-3 (1-1746) x US-10-451-168-91 (1-575)

```

```

QY 34 TGGTCCTTTTTCGTATTCACGCCCTGTGGACACAGACAGTGGGAAAACTGGG 93
Db 14 TrpAlaIleTrpValLeuAlaIaPro-----ThrlYsgly 25
QY 94 CTTTGTCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGATC 153
Db 26 ProSerAlaGUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGl 45
QY 154 ACTGTGCTGGGAAGCCGTGCTGTGAACGAGTTCCTGTGAGTCCCTTGTGCTCC 213
Db 46 ThrValLeuGlySerProAlaProValAsnValPheLeuGlyValaProPheAlaIaPro 65
QY 214 CCGTCGGATCCCTCGATTTACGAACCGCAGCCTGATCCCTGGGATTAATTGCGA 273
Db 66 ProleuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArg 85
QY 274 GAAGCCACTCTTACCTTAATTTGCTCTCAAGAACTCAAGATGCTGCTTTAGTCAA 333
Db 86 GUAAlaThrSerTyProAsnLeuGlyLeuGlnAsnSerGlnTrpLeuLeuAspGln 105
QY 334 CACATGCTCAAGTGCATTAACCCGAAATTCGGAAGTGCAGAGCTCCTTACCTGAAC 393
Db 106 HisMetLeuLysValHisTrpProLysPheGlyValaSerGluAspCysLeuTyLeuAsn 125
QY 394 ATCTATGCGCCTGCCACGACGATACAGGCTCCAGACTCCCGCTTGGTGTGTTCCCA 453
Db 126 IleTyAlaProAlaHisAlaAspThrGlySerTybLeuProValLeuValTrpPhePro 145
QY 454 GGAGGTGCTTCAAGATGAGCTCAGCTTCATCTTTGATGGTCCGCTGCTGCTACT 513
Db 146 GlyGlyAlaPheTyThrlGlySerAlaSerIlePheAspGlySerAlaLeuAlaIaTyx 165
QY 514 GAGGACGTGCTGGTGTGCTGCTCCAGTACCCGGCTNAGAAATTTGGTTCCTTCAACA 573
Db 166 GluAspValLeuValaValaValaGlnTyArgLeuGlyIlePheGlyPhePheThrThr 185
QY 574 TGGATCAGCATGCTCCGGGGAATCGGCTTCAAGACACAGTGGCTGCTGCTGG 633
Db 186 TrpAspGlnHisAlaProGlyAsnTrpAlaPheTybAspGlnValaAlaAlaLeuSerTrp 205
QY 634 GTCCAGAAAGACATGAGTCTTCGATGGGGAACCCAGCTGTGAGCAATCTTTGGCAG 693
Db 206 ValGlnLysAsnIleGlnPhePheGlyAspProSerSerValTrhIlePheGlyGln 225
QY 694 TCCGCGGGAAGCCATAAGTGTTCAGTCTTAATCTGTCTCCATGGCCAAAGCTTATTC 753
Db 226 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaYsglyLeuPhe 245
QY 754 CACAAAGCATATGAGAGAGTGGGGGACATCCATCCCTTACCTGGAGGCCATGATTAT 813
Db 246 HisLysAlaIleMetGlnSerIleValaIleIleProTyLeuGlnAlaHisAspTyx 265
QY 814 GAGAAGAGTGAAGACCTGAGGTGTGACATTTCTGTGTAGCAATGCTCAGACTCT 873
Db 266 GluLysSerSerGluAspLeuGlnValaValaHisPheCysGlyAsnAsnAlaSerAspSer 285
QY 874 GAGGCCCTGTGAGGTGCTTGAAGAACAAACCTTCCAAAGAGCTGTGACCTCAGCCAG 933
Db 286 GluAlaLeuLeuArgCysLeuArgThrlYsProSerTybGlnLeuLeuThrLeuSerGln 305
QY 934 AAAAGAAAGCTTTTACCTCGAGTGGTGAATGCTGCTTCTTCCTAATAGAGCTTAGAT 993
Db 306 LysTrhLysSerPheThrArgValaValaAspGlyAlaPhePheProAsnGlnProLeuAsp 325
QY 994 CTATTGTCTCAGAAAGACTTTAAAGCAATTCCTTCATCATCGAGTCAATTAACGACGAG 1053
Db 326 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValaAsnAsnHisGln 345

```



```

QY 751 TTCCAAAGCCATCATGAGAGAGTGGGAGCCATCATCCCTTAAGCCGATCATGAT 810
DB PhehlsyvalailemectluserglyvalaiailelleProtyrleucluhahsaer 331
QY 811 TATGAGAGAGTGAAGACCTGACAGGTGGTTCGACATTTTGTGGTAAACAATGCGTCAAC 870
DB Tyrgluysserlglyabpueglnvalaiaahlsphesgylasnaabnlaseerap 351
QY 871 TGTGAGGCGCTGCTGAGGAGCTGAGGAGCAAAACCTCCAAAGAGGTGTGACCCCTCAC 930
DB SergluabaleuclargCyseuclarglryrProserlysgluuclueuthrleuser 371
QY 931 CAGAAAACAAGCTCTTTCATCGAGTGTGATGAGTGTCTTTCTTAATAGACCTCTA 990
DB GlnlystrlyrserPhehthrargvalaiaabpgeglyalapherProabngluproleu 391
QY 991 GATCATTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGGAGTCAATTAACAC 1050
DB AspleuclusergllyabpueglysalalelProserlleleglyalabnbnhs 411
QY 1051 GAGTGTGCTTCCTGCTGCTTAAGAAAGAGGCTCCGATGATCTCAGTGGCTCCAAAG 1110
DB GlnCysglyPheleucluserProcluysgluabProgluvalleucluserlyserbnlys 431
QY 1111 TCCCTTGGCCCTCATCTGATTAACAAACATCTGACATCCCGCTCAGTATTGACACTT 1170
DB Serleuabaleucluserleuclenlnleuclenlleleuclenlleleuclenlleleu 451
QY 1171 GTGGTAATGAATTAATCTTCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTTG 1230
DB ValaiaabngluyrPhehlsaspyshlsSerleuthrGlulleargasperleuclen 471
QY 1231 GACTTGTCTGAGAGTGTCTTTGTGTCCTGCTGCTGATCATGACAGTGTGATGACAG 1290
DB AspleucluserglyabpueglysalalelProabnglvalaiaabpgeglyalapher 491
QY 1291 GATGCTGGGACACTGTCTACTTCTATGAGTTCGACCGGCTCAGTCTTGAAGAC 1350
DB AspalaglyalabProvallyrPhehthrargluphehtharglsrProclnlyserhegluabp 511
QY 1351 AGGAAGCCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTTGTGGTGGT 1410
DB ThrlyrProclalPhevallysalalelPhehlsaspyshlsSerleuthrGlullearg 531
QY 1411 GCCTTCTGAGAGGCGACATTTGTATGTTTCAAGAGGACCGAGAGAGAGAAAGTACTG 1470
DB AspleucluserglyabpueglysalalelPhehlsaspyshlsSerleuthrGlullearg 551
QY 1471 AGCCGGAAGATGATAATTAAGGAGTACTTGTGCTGCAACCGGGAATGCTAATGGGAG 1530
DB SerarglyrserleucluserlyrtrPalaatrPhehlsaspyshlsSerleuthrGlullearg 571
QY 1531 GACCTGTCTCTGAGGACCTTAATTAATGACTGAGACGATCACTCCAGCTGAGCTTGAAC 1590
DB AspleucluserlyrtrPalaatrPhehlsaspyshlsSerleuthrGlullearg 591
QY 1591 ATGAGCTTGAAGACGATCAAGAAACCGGAGGAGATTTTGAACGACGACCATCCCG 1650
DB MetserleucluserglyabpueglysalalelPhehlsaspyshlsSerleuthrGlullearg 611
QY 1651 CTGATCTGTCTGAGGACCTGAGGAGTCAAGGCTGCTGCTTCTTCTTAATCTTCTGCT 1710
DB LeuclleucluserlyrtrPalaatrPhehlsaspyshlsSerleuthrGlullearg 631
QY 1711 CTCTCAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1743
DB 632 LeuclleucluserlyrtrPalaatrPhehlsaspyshlsSerleuthrGlullearg 642

```

RESULT 6
 US-10-451-168-92
 ; Sequence 92, Application US/10451168
 ; Publication No. US20040091969A1
 ; GENERAL INFORMATION:

```

/ APPLICANT: SMITHKLINE BEECHAM CORPORATION
/ APPLICANT: SMITHKLINE BEECHAM P.1.C.
/ APPLICANT: GLAXO GROUP LIMITED
/ TITLE OF INVENTION: NOVEL COMPOUNDS
/ FILE REFERENCE: GP50039
/ CURRENT APPLICATION NUMBER: US/10/451,168
/ CURRENT FILING DATE: 2003-11-12
/ PRIOR APPLICATION NUMBER: PCT/US01/49232
/ PRIOR FILING DATE: 2000-12-17
/ PRIOR APPLICATION NUMBER: 60/256,710
/ PRIOR FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 60/257,048
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 60/260,482
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: 60/264,922
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/266,797
/ PRIOR FILING DATE: 2001-02-06
/ PRIOR APPLICATION NUMBER: 60/276,988
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/281,535
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/289,622
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 92
/ LENGTH: 581
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-451-168-92

Alignment Scores:
Pred. No.: 1,076-224 Length: 581
Score: 2607.50 Matches: 499
Percent Similarity: 92.67% Conservative: 7
Best Local Similarity: 91.39% Mismatches: 25
Query Match: 82.26% Indels: 15
DB: Gaps: 3

US-10-023-515-3 (1-1746) x US-10-451-168-92 (1-581)
QY 34 TGTGCTTTTCTGATTCCTCAGGCCCTGTTGGGACACAGACAGTGGGAAACTGGG 93
DB 14 TTPAAllettrPvalbueuAlaPro-----Thrlysgly 25
QY 94 CCTTGTGAGAGGCGACAGAGAAACACAGGCTGGATTCAGGCAAGCAAGTC 153
DB 26 ProserAlagluglyrProglmrghsnhrarglueglyrtrPlleclnlysglnval 45
QY 154 ACTGTGTGGAGAGCCCTGCTGTGAAAGTGTCTCGAGAGTCCCTTGTGCTGCC 213
DB 46 ThrValleuglySerProvalProvalabnvalPheleuglyValrProPheAlaPro 65
QY 214 CCGCTGGATCCCTCGATTAACGAAACCGGAGCTGATCGCCCTGGGATACTTGGGA 273
DB 66 ProleucluserlyrtrPalaatrPhehlsaspyshlsSerleuthrGlullearg 85
QY 274 GAAGCACTCTTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTGAATCA 333
DB 86 GluAlatThrSerlyrProabngluproleucluserlyrtrPalaatrPhehlsaspyshls 105
QY 334 CACATGCTCAGGAGTATTAACCGAAATTCGAGTGTCAAGAACTGCTTCACTGAAC 393
DB 106 HlsMetleucluserlyrtrPalaatrPhehlsaspyshlsSerleuthrGlullearg 125
QY 394 ATCTATGCGCTTGGCCAGCGGATACAGGCTCAGAGCTCCCGCTTGTGTGTCCCA 453
DB 126 lIetyrAlaProclalPhehlsaspyshlsSerleuthrGlullearg 145
QY 454 GGAGGCTTCAAGACTGCTCAGGCTCAGCTTCTTGAATGGGTCCGCTGCTGCTAT 513

```

Db 146 GlyIyAlaPheIyThrcIySerAlaSerIlePheAerGlySerAlaLeuAlaIaTyx 165
Qy 514 GAGACGCGCTGGTGTGGTCCGTCCAGTACCGGCTAAGAAATATTTGGTTTCTTCCACCA 573
Db 166 GluAerValIeValIaValIaGlnTyArGleuGlyIlePheGlyPhePheThrThr 185
Qy 574 TGGATCAGCATGCTCCGGGGAATGGGCTTCAAGACACAGGTGGCTGCTCTGG 633
Db 186 TrpAerGlnIleAlaProGlyAerThrAlaPheIyAerGlnValAlaAlaLeuSerTrp 205
Qy 634 GTCCAGAAAGACATGAGTCTTCGCTGGGGAACCCAGCTCTGTAGCCATCTTTGGCGAG 693
Db 206 ValGlnTyAerIleGlyPhePheGlyIyAerProSerSerValThrIlePheGlyGln 225
Qy 694 TCCGGGGGAGCCATAGTGTCTTAGTCTTATACGTCTCCCATGGCCCAAGGCTTATTC 753
Db 226 SerAlaGlyAlaIleSerValSerSerIleLeuSerProMetAlaIyGlyLeuPhe 245
Qy 754 CACAAAGCCATCATGAGAGTGGGAGGCGCATATCCCTTACCTGGAGGCCATGATTAT 813
Db 246 HisIyValaIleMetGlySerGlyAlaAlaIleIleProTyLeuGlnAlaHisAerTyx 265
Qy 814 GAGAAAGTGAAGACCTGAGGTGGTTCACATTTCTGTGTAAACATGCGTCAAGTCT 873
Db 266 GluTySerGlnAerPheGlnValValAlaAlaHisPheCyGlyAerAerAlaSerAer 285
Qy 874 GAGGCGCTGTAGGTGCTTGAAGACAAACCTCCAAAGACCTGTGACCTTCAAGCCAG 933
Db 286 GluAlaLeuLeuAerCyLeuAerGlyThrTyProSerTyGlyLeuLeuThrLeuSerGln 305
Qy 934 AAAACAAAGTCTTCACTCGAGTGTGATGCTGCTTTCTTCTAATAGAGCTTAGAT 993
Db 306 LysThrTySerPheThrAlaGlyAlaAerGlyAlaAerPheProAerGlnProLeuAer 325
Qy 994 CTATTGTCTCAGAAACATTTAAACCAATCTTCCATCATGCGAGTCAATACCAAGAG 1053
Db 326 LeuLeuSerGlnTyAlaPheTyAlaIleProSerIleIleGlyValAerAerHisGln 345
Qy 1054 TGTGGCTTCCGTGCTTGAAGAG-----GCTCCGAAATCCTCAGTGGCTCCAC 1107
Db 346 CyGlyPheLeuLeuProMetValArgIleLeuAlaValHisThrAlaThrProSerAer 365
Qy 1108 AAGTCCCTGGCCCTC-----CATCGATACAAACATCCGTGACATCCCG 1152
Db 366 ArgAerAlaAlaLeuAlaSerThrAlaGlyHisPheHisAerGlyHisGlnHisIlePro 385
Qy 1153 CCTCAGATTTGCACCTTGTGGCTAATGAATCTTCCATGACACACATCCCTGACTGA 1212
Db 386 ProGlnTyLeuHisLeuValAlaAerGlnTyPheHisAerPheHisSerLeuThrGln 405
Qy 1213 ATCCGACACAGTCTTGGACTTGTGGAAGATGTGTTCTTGTGTGCTCCCTGACTGATC 1272
Db 406 IleAerAerSerLeuLeuAerPheLeuGlyAerValPhePheValValProAlaLeuIle 425
Qy 1273 ACAGCTCATATCACAAGATCTGATGCACTGCTACTCTTATGAGATTTGCGACACGG 1332
Db 426 ThrAlaAerGlyHisAerAerAlaGlyAlaProValTyPheTyGlyPheAerGlyHisArg 445
Qy 1333 CCTCAGTCTTTGAAGACACGAGCCGCTTTGTCAAGCCGACACACGCTGATGAAGTC 1392
Db 446 ProGlnTyPheGlnAerThrTyProAlaPheValTyAlaAerPheHisAlaAerGlnVal 465
Qy 1393 CGCTTGTGTTCCGCTGGTCCCTTCTGAAGGGGACATTTGTTATGTTGGAAGAGACCG 1452
Db 466 ArgPheValPheGlyAlaPheLeuTyGlyAerPheValIleValMetPheGlnIyAlaThr 485
Qy 1453 GAGGAGGAGAGTCTGAGCCGGAAGATGATGAATCTGGGCTACCTTCTGGAAC 1512
Db 486 GluGlnGlyTySerLeuSerAerGlySerMetGlyTyTrpAlaThrPheAlaArgThr 505
Qy 1513 GGGATCTTAATGGGAACACCTGTCTGTGGCCAGCTTAATATCTGACTGAGCAGTAC 1572
Db 506 GlyAerProAerGlyAerAerPheSerLeuTrpProAlaTyAerHisLeuThrGlnGlnTyx 525

Qy 1573 CTCCAGCTGGACTTGAACATGAGCCTCGAGAGAGACTCAAGAACCGCGGCTGATTTT 1632
Db 526 LeuGlnLeuAerPheAerMetSerLeuGlyGlnArgLeuTyGlnProAlaArgAerVal 545
Qy 1633 TGGACGACACCATCCCC 1650
Db 546 TrpValThrGlyTyxPro 551
RESULT 7
US-10-094-749-2375
; Sequence 2375, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2375
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2375
Alignment Scores:
Pred. No.: 1,4e-224 Length: 525
Score: 2606.00 Matches: 503
Percent Similarity: 88.77% Conservative: 3
Best Local Similarity: 88.25% Mismatches: 6
Query Match: 82.21% Indels: 58
DB: 15 Gaps: 2
US-10-023-515-3 (1-1746) x US-10-094-749-2375 (1-525)
Qy 34 TGGTGTCTTTTCTCGATTCTCCAGCCCTGTGTGACACAGACAGTGGGAAAACTGGG 93
Db 14 TrpAlaIleTrpValLeuAlaIaPro-----ThrIyGly 25
Qy 94 CTTTCTGTGAAGGCCACAGAGAACCCAGCTGGATGATTCAGGGGACGAAGTC 153
Db 26 ProSerAlaGlnGlyProGlnAerAerThrArgLeuGlyTrpIleGlnGlyVal 45
Qy 154 ACTGNGCTGGAAAGCCCTGCTGTGAACGTGTTCTCGAGTCCCTTGTGCTGCTCC 213
Db 46 ThrValLeuGlySerProValProValAerValPheLeuGlyValProPheAlaIaPro 65
Qy 214 CCGCTGGATCCCTGCAATTAGAACCCGACGCTGATGCGCTTGATTAACCTTGGA 273
Db 66 ProLeuGlySerLeuAerGlyPheThrAerProGlnProAlaSerProTrpAerAerHisArg 85

QY 274 GAAGCACCTCTACCTTAATTGTCCTCCAGAACTCAGAGTGTGCTTATGATCA 333
 DB 86 GUAUATThSerTyPProbsnleucylsleuGlnsmSerGluTrpneuLeuamspGln 105
 QY 334 CACATGCTCAGAGTGCATTACCCGAAATTCGAGTGTCCAGAAAGATGCTCTTACCTGAAC 393
 DB 106 HsmetLeuylValHlstrProlysPheGlyValSerGluabpSerLeuTyRleuam 125
 QY 394 ATCTATGCCCTGCCACGCCGATACAGGCTCCAGACTCCCGCTCTTGTTGGTGTCCCA 453
 DB 126 ILeTyRAlaProAlaHsAlaapThrGlySerTyLeuProValLeuValTrpPhePro 145
 QY 454 GGAAGTGCCTTCAGAACTGAGCTCAGGCTCCATCTTGTATGGGTCGCGCTGAGTGCCTAT 513
 DB 146 GlyGlyAlaPheLeuTyRTrpGlySerAlaSerTlePheabpGlySerAlaLeuAlaAlaTyR 165
 QY 514 GAGACGTGCTGTGTGTGTGTCGTCCAGTACCGGCTACGAAATTTGGTTTCTTACACACA 573
 DB 166 GluabpValLeuValValValValGlnTyRArgLeuGlyIlePheGlyPhePheThrThr 185
 QY 574 TGGGATCAGCATGCTCCGGGGAATGGGCTTCAGAGACAGAGTGTGCTGTCTGTGG 633
 DB 186 TTPaapGlnHsAlaPProGlyAsnTrpAlaPheTyAspGlnValAlaAlaLeuSerTyP 205
 QY 634 GTCCAGAAAGACATGAGTCTTTCGSGTGGGACCCGAGCTGTGTACCATCTTTGGCGAG 693
 DB 206 ValGlnTyAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGln 225
 QY 694 TCCGCGGAGCATAGTGTCTTCTATGCTTATATCTGTCTCCATGGCCAAAGCTTATTC 753
 DB 226 SerAlaGlyAlaIleSerAlaSerSerLeuIleLeuSerProMetAlaIlysglyLeuPhe 245
 QY 754 CACAAAGCCATATGAGAGAGTGGGCTGACCATCATCTTACCTGGAGGCCATGATTAT 813
 DB 246 HsTyRAlaIleMetGlnSerGlyValAlaIleIleProTyRLeuGlnAlaHsAspTyR 265
 QY 814 GAGAAGAGTGAAGACCTGAGGCTGTGACATTTCTGTGTAAACAATCGTCAGACTCT 873
 DB 266 GlnTyRSerGlnAspLeuGlnValAlaAlaHsPheCysGlyAsnAsnAlaSerAspSer 285
 QY 874 GAGGCTCTGCTGAGTGTGCTGAGACAAACCTCCAGAGAGTGTGACCTCTCAGCCAG 933
 DB 286 GluAlaLeuLeuArgCysLeuArgThrTyRProSerTyRLeuLeuLeuThrLeuSerGln 305
 QY 934 AAAACAAAGCTTTTCACTGAGTGTGTGATGGTCTTTCTTCTTATATGAGCTCTTATAG 993
 DB 306 TyRThrTyRSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 325
 QY 994 CTAATGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATATGGAGTCAATAACCCAG 1053
 DB 326 LeuLeuSerGlnTyRAlaPheTyRAlaIleProSerTleIleGlyValAsnAsnHsGln 345
 QY 1054 TGTGGCTTCTGCTGCTATGAGAGAGGCTCCAGATCCTCAGTGGCTCCAAAGATGCC 1113
 DB 346 CysGlyPheLeuLeuProMetTyRsglValaProGlnIleLeuSerGlySerHsTyRAsp 365
 QY 1114 CTTGGCTTCTCATGTATACAAACATCTGCACATCCGCTCATGATTTTGGACCTTGTG 1173
 DB 366 LeuAlaLeuHsIleuIleGlnAsnIleLeuHsIleProProGlnTyRLeuHsIleVal 385
 QY 1174 GCTATGAAATACTTCCATGACAGACACTCCCTGACCTGAATTCGAGACAGTCTTGGAC 1233
 DB 386 AlaAsnGlnTyRPhetHsAspTyRHsSerLeuThrGlnIleArgAspSerLeuLeuAsp 405
 QY 1234 TTGCTTGAGATGTGTCTTGTGTGCTGCTGACATGATCAGAGTGTGATATCAGAGAT 1293
 DB 406 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyRHisArg--- 424
 QY 1294 GCTGTGCACTGTACTTCTATGAGTTCGGCACCGGCTCAGTGTCTTGAAGACAG 1353
 DB 424 ----- 424

QY 1354 AAGCCGGCTTTTGTGTAAGCCGACCAAGCTGATGAAGTCCGTTGTGTGCGTGTGCC 1413
 DB 424 ----- 424
 QY 1414 TTCTCGAAGGGGACATTGTATTGTTCCAGAGAGCCACGAGAGAGAAATTACTGAC 1473
 DB 425 -----GluGlyAlaThrGluGluGluTyRLeuLeuSer 435
 QY 1474 CGAAGATGATGAATTACTGGGCTACCTTGTGCTCGAACCGGGAATCTTAATGGGAACGAC 1533
 DB 436 ArgTyRMetMetLeuTyRTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 455
 QY 1534 CTGTCTGTGTGCGCAGCTTATATCTGACTGACAGTACCTTCAAGTGTGATGAACATG 1593
 DB 456 LeuSerLeuTrpProAlaTyRAsnLeuThrGlnGlnTyRLeuGlnLeuAspLeuMet 475
 QY 1594 AGCTTCGACAGAGACTCAAGAAACCGGGGTGATTTTGGACCAAGCACTCCCTCTG 1653
 DB 476 SerLeuGlyGlnArgLeuTyRsglProArgValaAspPheTrpHsSerThrIleProLeu 495
 QY 1654 ATCTGTGCTGCTCCGACATGCTCCACAGTCTCTTCTTCTTAACTTCCCTGTCTC 1713
 DB 496 IleLeuSerAlaSerAspMetLeuHsSerProLeuSerSerLeuThrPheLeuSerLeu 515
 QY 1714 CTCCAGCTTCTTTTCTTTTGTGCTCTCT 1743
 DB 516 LeuGlnProPhePhePheCysAlaPro 525
 RESULT 8
 US-10-114-270-196
 ; Sequence 196, Application US/10114270
 ; Publication No. US20040030110A1
 GENERAL INFORMATION:
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Malvankar, Uriel M.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Patursajan, Meera
 ; APPLICANT: Liu, Zhaozhong
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Li, Li
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Zernusen, Bryan D.
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Pena, Carol E.A.
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Burgess, Catherine E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Padigar, Muralidhara
 ; APPLICANT: Shimkete, Richard A.
 ; APPLICANT: Gangoli, Bsha A.
 ; APPLICANT: Taupier Jr., Raymond J.
 ; APPLICANT: Casman, Stracie J.
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Lieke, Mario W.
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Rothenberg, Mark B.
 TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same
 FILE REFERENCE: 21402-322C
 CURRENT APPLICATION NUMBER: US/10/114,270
 PRIORITY FILING DATE: 2002-11-27
 PRIOR APPLICATION NUMBER: 60/281,086
 PRIOR FILING DATE: 2001-04-03
 PRIOR APPLICATION NUMBER: 60/281,136
 PRIOR FILING DATE: 2001-04-03
 PRIOR APPLICATION NUMBER: 60/281,863
 PRIOR FILING DATE: 2001-04-05


```

; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 196
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-196

Alignment Scores:
Pred. No.:      8,476-224      Length:      581
Score:          2597.50       Matches:     497
Percent Similarity: 95.07%    Conservative: 4
Best Local Similarity: 94.31%  Mismatches:  17
Query Match:    81.94%       Indels:      9
DB:             15           Gaps:        3

US-10-023-515-3 (1-1746) x US-10-114-270-196 (1-581)
QY 91 GGGCCTTGTGCTGAAGGGGCGACAGAGGAACACAGGCTGGATGATTCAGGGCAAGCA 150
DB 27 GYPRoserAlaGluGlyProGlnArgAsnTrpArgLeuGlyTrpLleGlnGlySerGln 46
QY 151 GTCACTGTGCTGGGAAGCCTGTGCTGTGAACGTTCCTCGAGTCCCTTTGTGCTGT 210
DB 47 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIle 66
QY 211 CCCCCGTGGGATCCCTCGGATTTACGAACCCGACGCTGCATCGGCTGGGATTAATTG 270
DB 67 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 86
QY 271 CGAGAGGCACTCCTACCTAATTTGTGCTCCAGAACTCAGAGGCTGCTCTTAGAT 330
DB 87 ArgGluAlaThrSerTrpProAsnLeuCysLeuGlnAsnSerGlnTrpLeuLeuAsp 106
QY 331 CAACACATGCTCAAGAGTGCAATTAACCGAAATTCGAGTGTCAAGAGACTGCTTACCTG 390
DB 107 GlnHisMetLeuLeuValHisTrpProLysPheGlyValSerGlnAspCysLeuTrpLeu 126
QY 391 AACATCTATGCGCTGCGCCACGCGCGATACAGGCTCCAGCTCCCGTCTTGTTGTTTC 450
DB 127 AsnIleTrpAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhe 146
QY 451 CCAGAGAGGCTTGAAGCTGCTCAGCTCAGCTCATCTTTGAGGGGCTGCGGCTGCTGC 510
DB 147 ProGlyGlyAlaPheLysThrGlySerAlaSerLlePheAspGlySerAlaLeuAlaIle 166
QY 511 TATGAGAGCTGCTGTTGTGTGCTGCAGTACCGGCTAGGAATTTGGTTTCTTACCC 570
DB 167 TyrGluAspValLeuValValValGlnTrpArgLeuGlyLlePheGlyPhePheThr 186
QY 571 ACATGGGATCAGCATGCTCCGGGGAACCTGGGCTTCAAGAACCAAGTGGCTGCTGTGC 630
DB 187 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaIleAsnSer 206
QY 631 TGGGTCAGAGAGAACATGAGTTCTTCGGTGGGGAACCCAGACTCTGTGACATCTTTGGC 690
DB 207 TrpValGlnLysAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly 226
QY 691 GAGTCGCGGGAGCCATAGTGTCTTACTTATAGTGTCTCCATGGCCAAAGGCTTA 750
```

```

DB 227 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 246
QY 751 TTCCCAAGCCATATATGAGAGTGGGTGGCCATCATCCCTTACCTGAGGCCCATGAT 810
DB 247 PheHisLysValAlleMetGluSerLysValAlaIleIleProGlyLeuGlnAlaHisAsp 266
QY 811 TATGAGAGAGTGAAGCACTCAGAGTGGTGCACATTTCTGTGTGAACAATCGTCAAG 870
DB 267 TyrGlnLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnHisAsp 286
QY 871 TCTGAGCCCTGCTGAGGTGCTGAGCAAAACCCCTCAAGAGACTGTGACCTGACG 930
DB 287 SerGlnAlaLeuLeuArgCysLeuArgThrLysProSerLysGlnLeuLeuThrLeuSer 306
QY 931 CAGAAAACAAAGTCTTTCACTCGAGTGTGATGAGTCTTTCTTCTTAATGAGCTCTTA 990
DB 307 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGlnProLeu 326
QY 991 GATCTATTGTCTCAGAAAGCAATTTAAACAATTCCTTCATCATCGAGTCAATTAACAC 1050
DB 327 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHis 346
QY 1051 GAGTGTGCTTCTGCTGCTGCTATGAAGAG-----GCTCCTGAGATCTCAGTGCCTCC 1104
DB 347 GluCysGlyPheLeuLeuProMetValArgLleLeuAlaValHisThrAlaThrProSer 366
QY 1105 AACAAAGTCCCTTGCCCTC-----CATCGATACAAACAAATCCTGCACATC 1149
DB 367 AsnArgAspAlaAlaLeuAlaSerThrAlaGlyLysPheHisAspArg-----HisLle 384
QY 1150 CCGCTCAGATATTTGACCTTGAGCTAATGAAATCTTCCATGACCAAGACATCCCTGACT 1209
DB 385 ProProGlnTrpLeuHisLeuValAlaAsnGlnTrpPheHisAspLysHisSerLeuThr 404
QY 1210 GAAATCCGAGACAGTCTTCTGAGCTGCTTGAAGATGTGTTCTTTGTGCTCCCTGACTG 1269
DB 405 GluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValValProAlaLeu 424
QY 1270 ATCAACAGCTCAATATCAAGAGATGCTGTGCACTGTCTTCTTAATGAGTTCCGGAC 1329
DB 425 IleThrAlaArgLysTrpHisAspArgAlaGlyAlaProValTrpPheTrpGlnPheArgHis 444
QY 1330 CCGCTCAGTCTTGAAGACACGAAAGCGGCTTTGTCAAGCGGACCAAGCTGTGATGA 1389
DB 445 ArgProGlnCysPheGluAspThrTrpProAlaPheValLysAlaAspHisAlaAspGln 464
QY 1390 GTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGACATTTGTTATGTTGGAAGGAGCC 1449
DB 465 ValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlnGlyAla 484
QY 1450 ACGGAGAGGAGAGATTACTGACCGCGAAGATGAAATCTGGGGCTACCTTGCTGCTGA 1509
DB 485 ThrGlnGlnGlnLysLeuLeuSerArgLysMetCysLysTrpTrpAlaThrPheHisAlaArg 504
QY 1510 ACCGGGAATCTTAATGAGAGACCTGTCTGTGGCCAGCTTAAATCTGACTGAGCAG 1569
DB 505 ThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeuThrGlnGln 524
QY 1570 TACCTCCAGCTGGAATTGAACATGAGCTTGGACAGAGACTCAAGAACCCGGGCTGAT 1629
DB 525 TyrLeuGlnLysAspLeuAsnMetSerLeuGlnArgLeuLysGlnProArgArgAsp 544
QY 1630 TTTTGGACCAAGACCATCCCC 1650
DB 545 ValTrpValThrGlyTyrPro 551

RESULT 9
US-10-381-898-2
; Sequence 2, Application US/10381898
; Publication No. US2004008687A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valde, BAUGHN, Mariah R.;
```


APPLICANT: BOROMSKI, Mark L.; DING, Li;
APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
APPLICANT: HAPPAIA, April J A.; ISON, Craig H.;
APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
APPLICANT: RAMKUMAR, Jayalakmi; RING, Huijun Z.;
APPLICANT: SANJANMALA, Madhusudan M.;
APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
APPLICANT: CHAWLA, Nandinder K.; WALSH, Roderick T.;
APPLICANT: WARREN, Bridget; XU, Yuming;
APPLICANT: YANG, Junming; YAO, Monique; YUE, Henry
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-023 USN
CURRENT APPLICATION NUMBER: US/10/381, 898
CURRENT FILING DATE: 2003-10-17
PRIOR APPLICATION NUMBER: PCT/US01/30662
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/236, 947
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/238, 864
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/242, 323
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/247, 581
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249, 519
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252, 834
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/250, 567
PRIOR FILING DATE: 2000-11-30
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7473645CD1
US-10-381-898-2

Alignment Scores:
Pred. No.: 5,596-223 Length: 618
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 81.66% Indels: 23
DB: 15 Gaps: 1

US-10-023-515-3 (1-1746) x US-10-381-898-2 (1-618)
QY 91 GGGCTTTCGCGAAGGCGACAGAGAAACACAGGCTGGATTCAGAGGCAAGCA 150
DB 92 GtlyProserAlaGlnGlyProGlnArgAsnThrArgLeuGlyTyrIleGlnIlyysGln 111
QY 151 GTCACTGTGCTGGAGAGCCCTGTGCTGCGAAGCTGTTCTCGAGTCCCTTGTGCT 210
DB 112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 131
QY 211 CCCCCGTGGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 270
DB 132 ProPheLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 151
QY 271 CGAAGAACCACTCTCTACCTTAATTTGCTTCGCGAAGCTCAGAGTGGCTCTTAAGT 330
DB 152 ArgGlnAlaThrSerTyrProAsnLeuCySleuGlnAsnSerGluTyrLeuLeuAsp 171
QY 331 CAACACATGCTCAGAGTGATTAACCGAATTCGAGAGTGCAGAGAGACTGCTTACCTG 390
DB 172 GlnHisMetLeuIlyValHisTyrProIlyPheGlyValSerGluAspCySleuTyrLeu 191

QY 391 AACATCTATGCGCTGCCACGCCGATACAGGCTCCAAAGCTCCCGTCTTGCTGCTTC 450
DB 192 AsnIleTyrAlaProAlaHisIleAspThrGlySerIlySleuProValLeuValTrpPhe 211
QY 451 CGAAGAGGCTCCTTCAAGACTGGCTCAGACCTTCATCTTTGATGGTCCGCTGCTGCC 510
DB 212 ProGlyGlyAlaPheIlyThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 231
QY 511 TATGAGACGTGCTGCTGCTGCTGCTCAGTACCGGCTAGAGAAATTTGGTCTTTCACC 570
DB 232 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 251
QY 571 ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGAACAGAGTGGCTCTGCTCC 630
DB 252 ThrTrpAspGlnHisAlaIleProGlyAsnThrAlaPheIlyAspGlnValAlaAlaLeuSer 271
QY 631 TGGGTCCAGAAAGACATCGAGTTCTTCGATGGGAAACCCAGCTGTGACCATCTTGGC 690
DB 272 TrpValGlnIlyAsnIleGluPhePheGlyIlyAspProSerValThrIlePheGly 291
QY 691 GAGTCCGGGGAGCCATAAGTGTTCCTTAATCTTATACGTCTCCATGGCCAAAGCTTA 750
DB 292 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIlySlyLeu 311
QY 751 TTCACAAAGCATGATGAGAGTGGGGGCGGCATCATCCCTTACCTGGAGGCCCATGAT 810
DB 312 PheHisIlyValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 331
QY 811 TATGAGAAAGTGAAGACCTGAGGTGGTTGACATTTCTGTGTACATAGCTGCAGAC 870
DB 332 TyrGluIlySerGluAspLeuGlnValValAlaHisPheCySlyAlaAsnAsnIleAsp 351
QY 871 TTTGAGGCTCTGCTGAGGTGCTTGAAGACAAACCTCCAGAGAGCTGTACCTCAGC 930
DB 352 SerGlnAlaLeuLeuArgCySleuArgThrIlySerProSerIlySlyIleuLeuThrLeuSer 371
QY 931 CAGAAAGCAAGTCTTTCATCTCGAGTGGTGAATGGGCTTCTTCCTAATAGGCTCA 990
DB 372 GlnIlySerThrIlySerPheThrArgValAlaAspGlyAlaAspPheProAsnIlyProLeu 391
QY 991 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATGAGAGTCAATACAC 1050
DB 392 AspLeuLeuSerGlnIlyAlaPheIlySalIleProSerIleIleGlyValAsnAsnHis 411
QY 1051 GAGTGTGCTTCTGCTGCTGCTATGAAAGAGGCTCCTGAGATCTAGTGGCTTCAAAG 1110
DB 412 GluCySlyIlyPheLeuLeuProMet ----- 419
QY 1111 TCCCTTGCCCTTCATCTGATACAAACATCTCGACATTCGCCCTCAGATTTGGACCTT 1170
DB 420 -----HisIleProProGlnIlyThrHisIleu 428
QY 1171 GNGGTAATGAATTACTTCATGACAGACACTCCGACTGAAATCGAGACAGTCTTCTG 1230
DB 429 ValAlaAsnGluTyrPheHisAspIlyShiSerLeuThrGlnIleArgAspSerLeuLeu 448
QY 1231 GACTTGTCTTGAGATGTGTCTTGTGCTGCTGCTGCTGATCAGAGCTGATATACAGA 1290
DB 449 AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArg 468
QY 1291 GATGCTGGTGCACCTGTCTACTTCTATAGATTCCGACCGGCTCAGAGCTTTGAAGC 1350
DB 469 AspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCyAspPheGluAsp 488
QY 1351 ACGAAGCGGCTTTGTGTCAAGCGCAGACCGCTGATGAAGTCCGCTTGTGTTCGATGT 1410
DB 488 ThrIlySerProAlaPheValIlyValAlaAspHisAlaAspGlyValAlaArgPheValPheGlyGly 508
QY 1411 GCTTCTCGAAGGAGGACATTTGTATGTTTGAAGAGCCACGAGAGAGAGAAAGTTACTG 1470
DB 509 AlaPheLeuIlyGlyAlaAspIleValIleMetPheGluGlyAlaThrGluGluGluIlySleuLeu 528

US-10-233-933A-4

Alignment Scores:

Pred. No.:	8.17e-169	Length:	542
Score:	1985.00	Matches:	373
Percent Similarity:	83.27%	Conservative:	60
Best Local Similarity:	71.73%	Mismatches:	85
Query Match:	62.62%	Indels:	2
DB:	16	Gaps:	2

US-10-023-515-3 (1-1746) x US-10-233-933A-4 (1-542)

```
OY 91 GGGCTTGTGTGAAGGGCCAGAGAAACACAGAGCTGGATGATTCAGGGCAAGCA 150
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 22 GlycerolAalAAsPAlaProValArgSerThrArgLeuGlyTrrValArgGlyysGln 41
OY 151 GTCACTGTGTGGGAAGCCCTGTGCCTGTGAACGTGTTCCTGGAGTCCCTTGTGCT 210
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 42 ThrThrValLeuGlySerThrValProValAAsnMetPheLeuGlyIleProTyrrAlaAla 61
OY 211 CCCCCGCTGGATCCCTGGATTTAGAACCCGAGCCCTGCATCCCGCTGGATTAAGTTG 270
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 62 ProProLeuGlyrProLeuArgPheLeuGlnProTyrrProAlaLeuProGlyAAsnAspPhe 81
OY 271 CGAGAAAGCCACCTCTACCTTAATTGTGTGCTCCAGAACTCAGAGTGGCTGCTTAGAT 330
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 82 ArgAAsnAlaThrSerTyrrProTyrrLeuGlyrPheGlnAAspLeuGlnTrrPheuValSerTyrr 101
OY 331 CAACACATGTCTCAAGGTGCATTACCCGAAATTCGAGGTGTCAAAAGATGTGCTTACCTG 390
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 102 GlnHlsValAlaLeuValAlaGlyrrProTyrrLeuGlnAlaSerGlnAAspTyrrLeu 121
OY 391 AACATCTATGGGCTGGCCGAGCGCATACAGGCTCCAGGCTCCCGCTGTGTGGTGTTC 450
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 122 AAsnIleTyrrAlaProAlaHlsAAsnAsnGlySerAAsnLeuProAlaMetValIrrPhe 141
OY 451 CCAGAGAGTGCCTTCAAGACTGCTCAGGCTCATCTTGTGATGGGTCCGCTGGCTGCC 510
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 142 ProGlyGlyAlaPheLeuMetGlySerAlaSerSerPheAAspGlySerAlaLeuAlaAla 161
OY 511 TATGAGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 162 TyrrGlnAAspValLeuIleValIrrThrGlnTyrrArgGlyGlyIlePheGlyPheAAsp 181
OY 571 ACATGGGATCAGAGCTCCCGGGGAACCTGGGCTTCAAGGACAGAGTGGCTGTGTGCC 630
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 182 ThrGlyAAspGlnHlsAAsnGlyAAsnTrrAlaLeuLeuAAspGlnValAlaAlaLeuThr 201
OY 631 TGGGTCCAGAAAGATCAGATGTTCTTCGATGGGGAACCCAGCTCTGTGACATCTTTGGC 690
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 202 TrrValAAsnAAsnAAsnIleGlnPhePheGlyGlyAAspProAAspSerValIrrIlePheGly 221
OY 691 GAGTCCGGGGAAGCCATTAAGTGTTCATGCTTATATCTGTCTCCATGGCCAAAGCTTA 750
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 222 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAAsnGlyLeu 241
OY 751 TTCCCAAGGATCATGAGAGAGTGGGAGGCTCATCTCCCTTAAGCTG---GAGGCCAT 807
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 242 PheHlsValAlaIleMetGlnSerGlyValAlaIleLeuProLeuLeuMetArgProPro 261
OY 808 GATTATGAGAAGTGAAGACCTGAGAGTGTGTGACATTTCTGTGTGAACATGCTGCA 867
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 262 GlyAAspGlnAAsnGlyAAsnAAsnLeuAlaArgIleCyGlyGlyHlsAAsnAAsn 281
OY 868 GACTGTGAGGCTCTGTGAGTGTGTGAGGACAAACCTCCAGAGAGTGTGTGACCTTC 927
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 282 AAspSerAlaAlaLeuLeuGlnGlyAAsnAAsnAAsnAAsnAAsnAAsnAAsnAAsn 301
OY 928 AGCGAAGAAACAAAGCTTTCATCTGAGAGTGTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGT 987
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 302 SerTyrrValLeuThrPheSerIleProValIleAAsnAAsnAAsnAAsnAAsnAAsn 321
OY 988 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATGATCGAGTCAATAAC 1047
```

```
DB 322 ValAlaLeuLeuThrGlnIleAlaPheAAsnSerValProSerIleIleGlyValAAsnHsn 341
OY 1048 CACGAGTGTGCTTCTCTGCTGCTATGAGAAGGCTCTGTGAGTCTTCAAGTCTCCAAC 1107
DB 342 HlsGlnCyAAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyGlySerHsn 360
OY 1108 AAGTCCCTTGGCCCTCATCTGATACAAACATCCGAGCATCCGCTCAGATTAATTCGAC 1167
DB 361 ArgSerLeuAlaLeuTyrrLeuValHlsThrPheLeuAAsnIleProThrGlnTyrrHls 380
OY 1168 CTGTGCTATGAATATCTTCATGACAAAGCATCTCCCTGACTGAATTCGAGACAGTCTT 1227
DB 391 LeuValAlaAAspHlsTyrrPheTyrrAAsnHlsSerProValGlnIleAAspSerPhe 400
OY 1228 CTGACATGTCTTGGAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1287
DB 401 LeuAAspLeuLeuGlyAAspValLeuPheValAlaProGlyValIrrValThrAlaArgTyrrHls 420
OY 1288 AGAGATGCTGGAGACCTGTCTACTTCTATGAGTTTGGGACCGGCTCAGTGTCTTGA 1347
DB 421 ArgAAsnAlaGlyAlaProValIrrPheTyrrGlnPheGlnHlsProProGlnCyAAsn 440
OY 1348 GACACGAAGCCGCTTTTGTCAAAAGCCAGCACGCTGATGAAGTCCGCTTGTGTGCT 1407
DB 441 AAspThrArgProAlaPheValIrrPheTyrrGlnPheGlnHlsProProGlnCyAAsn 460
OY 1408 GTGTGCTTCTCGAAGGGGAGCATTTGTATGTTCCAAAGAGCCAGGAGAGAGAAAGTTA 1467
DB 461 GlyAlaPheLeuTyrrGlyAAspIleValMetPheGlnGlyAlaIrrGlnGlnTyrrLeu 480
OY 1468 CTGAGCCGGAAGATGATGAATATCTGAGGCTACTTGTGTGCAACCGGGAATCTTAAGG 1527
DB 481 LeuSerArgGlySmetMetAArgTyrrTrrAlaAAsnPheAlaAAsnGlnGlyAAspProAAsnIly 500
OY 1528 AACGACCTGTCTGTGTGCGACGCTTATATCTGACTGACAGTACCTTCAGCTGAGCTTG 1587
DB 501 GlnGlyValProLeuTrrProAlaTyrrThrGlnSerGlnGlnTyrrLeuTyrrLeuAAspLeu 520
OY 1588 AACATGAGCTTCGAGACAGAGACTCAAAAGACCGGCGGTGATTTTGGACACAGACCTTC 1647
DB 521 SerValSerValGlyGlnTyrrLeuTyrrLeuGlnGlnValGlnPheTyrrMetAAsnThrIle 540

RESULT 13
US-10-451-168-93
; Sequence 93, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FIDE REFERENCE: GPe50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
```

NUMBER OF SEQ ID NOS: 110
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 93
 LENGTH: 356
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-451-168-93

Alignment Scores:

Pred. No.: 1 74e-150 Length: 356
 Score: 1780.00 Matches: 340
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 56.15% Indels: 0
 DB: 15 Gaps: 0

US-10-023-515-3 (1-1746) x US-10-451-168-93 (1-356)

```

QY 724 ATACTGCTCCAGGCGCAAGGCTTATTCACAAAGCATGATGAGAGTGGGCTGCC 783
DB 17 lIeLuserProMeCaLaIaYsGlyLeuPheHlSbYsAlaIlEmeCsluserGlyValaLa 36
QY 784 ATCATCCCTTACCTGGAGGCCCAATGATTATGAGAGAGTGAAGACCTTGACGCTGTTGCA 843
DB 37 lIeLlPrlOtyrLeuGlulAhlIsaPtyrGlyLysSerGluApsLeuGlulValaLa 56
QY 844 CATTCTGTGTTAACAATGCGTCAGACTCTGAGGCGCTGAGGCTGAGGAGCAAAA 903
DB 57 HlSpheCyGlyYAsnAsnAlaSerApsSerGluAlaLeuAlyGlyLeuAlyrGlnLys 76
QY 904 CCCTCAAGAGAGCTGCTGACCTCAGCCAGCAAAAACAAGTCTTTCAGAGTGTGAT 963
DB 77 ProSerLysGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrValValaAps 96
QY 964 GGTGCTTCTTCTTAATGAGCCTTAATGATTATGTTCTCAAGAACATTAAAGCAATT 1023
DB 97 GlYAlaPhePheProApsGlnProLeuApsLeuSerGlnLysAlaPheLysAlaLe 116
QY 1024 CCTTCATATATGGAGTCAATAACCAAGATGGCTTCTGCTGCTATGAAAGAGGCT 1083
DB 117 ProSerLlelGlyYAlaAsnAsnHlSgLyGlyPheLeuLeuProMetLysGluAla 136
QY 1084 CCTGAGATCCTCAGTGGCTCAACAAGTCCCTGCTCATCTGATACAAACATCTG 1143
DB 137 ProGlnLleuSerGlySerAsnLysSerLeuAlaLeuHlSleuLleuGlnHlSleu 156
QY 1144 CACATCCCGCTCAGATTATGCACTTGTGGCTAATGAATACTTCATGACAGCACTCC 1203
DB 157 HlSlelProProGlnLtyrLeuHlSleuValAlaAsnGlnLtyrPheHlSAspLysHlSser 176
QY 1204 CTGACTGAATCCGAGACAGTCTTCTGACCTTGTGAGATGTGTTCTTTGTGGTCCCT 1263
DB 177 LeuThrGlnLlAArgApsSerLeuLeuApsLeuLeuGlyYAspValApsPheValAlaPro 196
QY 1264 GAGCTGATCAGCTCGATATACAGAGATGCTGGGCACTGCTCACTTCACTTATGAGTTT 1323
DB 197 AlaLeuLlelThrAlaArgLyrHlSArgApsAlaGlyAlaProValLyrPheLyrGlnPhe 216
QY 1324 CGGACCGGCTCAGTGTCTTGAAGAACAAGACCGGCTTGTCTCAAGCCGACACAGCT 1383
DB 217 ArgHlSArgProGlnLtyrPheGlnApsThrLysProAlaPheValLysAlaApsHlSAla 236
QY 1384 GATGAAGTCCGCTTGTGTTCCGTTGCTGCTTCTGAAGGGGACATGTTATGTTGAA 1443
DB 237 ApsGlnValAArgPheValApsGlyGlyAlaPheLeuLysGlyYAspLleValaMetPheGln 256
QY 1444 GAGGACGAGGAGAGAGAGTACTGAGCGGGAAGATGATGAATAACTGGGCTACTCTT 1503
DB 257 GlYAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetLysLyrTrpAlaThrPhe 276
QY 1504 GCTCGAACCGGGAATCTATGGGAACGACTGCTCTGTGGCCAGCTTAAATCTGACT 1563
DB 277 AlAArgThrGlyYAsnProApsGlnYAsnApsLeuSerLeuTrpProAlaLyrAsnLeuThr 296

```

```

QY 1564 GAGCAGTACCTCCAGCTGATCTTGAACATGACCTTCGACAGAGACTCAAGAACCGCGG 1623
DB 297 GluGlnLyrLeuGlnLeuApsLeuAsnMetSerLeuGlyGlnAArgLeuLysGlnProAArg 316
QY 1624 GTGATTTTGGACCAAGACCAATCCCTGATCTGTCTGCTCCGACATGCTCCCACT 1683
DB 317 ValApsPheTrpThrSerThrLleProLeuLleuSerAlaSerApsMetLeuHlSser 336
QY 1684 CCTCTTCTCTTACTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1743
DB 337 ProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAlaPro 356

```

RESULT 14

US-09-925-298-689
 Sequence 689, Application US/09925298
 Publication No. US20020039764A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA103
 CURRENT APPLICATION NUMBER: US/09/925,298

PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05881

PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270

NUMBER OF SEQ ID NOS: 846
 SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 689

LENGTH: 549

TYPE: PRT

ORGANISM: Homo sapiens

NAME/KEY: SITE

LOCATION: (1)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

LOCATION: (7)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-298-689

Alignment Scores:
 Pred. No.: 5 09e-101 Length: 549
 Score: 1230.00 Matches: 248
 Percent Similarity: 63.84% Conservative: 91
 Best Local Similarity: 46.70% Mismatches: 168
 Query Match: 38.80% Indels: 24
 DB: 9 Gaps: 8

US-10-023-515-3 (1-1746) x US-09-925-298-689 (1-549)

```

QY 91 GGGCTTCTGCTGAAGGCGCACAGAGAACACGAGTGGATTCAGGGCAAGCA 150
DB 16 GlYGlnApsSerAlaSerProLleArgThrThrHlSThrGlyGlnValLeuGlySerLeu 35
QY 151 GTCACGTCTCTGGAAAGCCCTGCTGCAACGTTCTCTGGAGTCCCCCTTGTGCT 210
DB 36 ValHlSValLysGlyYAlaAsnAlaGlyValGlnThrPheLeuGlyLleProPheAlaLys 55
QY 211 CCCCCGCTGGATCCCTGCGATTTCGAACCCGACGCTGCAATCGCCCTGGGTAACCTG 270
DB 56 ProProLeuGlyLProLeuAryPheAlaProProGlnLtyrProGlnSerLtyrPheSerGlyVal 75
QY 271 CGAAGACCACTCTCACTCACTTATTTGCTCTCAGAAC-----TCAGAG 315
DB 76 ArgApsGlyThrThrHlSProAlaMetCysLeuGlnApsLeuThrAlaValGlnSerGln 95
QY 316 TGGCTCTTATGATCAACAATGCTCAAGGTGCATTACCGCAAAATTCGAGTGCAGAA 375
DB 96 PheLeu-----SerGlnPheAsnMetThrPheProSerApsSerMetSerGln 111
QY 376 GACTGCTCTACCTGAACATATATGAGCTGCGCCAGCGCATACAGGCTCCAAAGTCCCC 435

```

Dd	112	AspCyLeuIyLeuSerIleIyThrProIaHisSerHisGluIySerHisLeuPro	131
Qy	436	GTCTTGCTGTGGTGTCCAGAGAGTGCCTTCAAGACTGCTCAGCCTCCATCTGATAGG	495
Dd	132	ValMetValTIpIleHisGlyIalaleuValPheGlyMetAlaSerLeuIyThrArgIy	151
Qy	436	TCGCGCCCTGGCTCCATAGAGAGCTGCTGTGGTGTGTGCTGCAGTACCGGCTAGAA	555
Dd	152	SerMetLeuAlaAlaLeuGlnIuHisValIalValIalIleIleGlnIyThrGluIyIal	171
Qy	556	TTTGGTTCTTTCACCAATGGGATAGCAATGCTCCGGGAACTGGGCGCTTCAAGACG	615
Dd	172	LeuGlyPhePheSerThrGlyAspIyHisAlaThrGlyAsnTrpGlyTyIyLeuAspGln	191
Qy	616	GTGGCTGCTGTCTCCGTGGGTCCAGAAAGAACATCGAGTTCTTGGTGGGAGACCCGCTCT	675
Dd	192	ValAlaAlaLeuAlaArgTIpValGlnGlnHisIleAlaHisPheGlyGlyAsnProAspArg	211
Qy	676	GTGACCACTTTTGGGAGTCCGGGGAGCCATAGTGTTTCTAGTCTTAACTGTCTCC	735
Dd	212	ValThrIlePheGlyGlySerAlaGlyIyThrSerValSerLeuValIalAsnPro	231
Qy	736	ATGGCCAAAGGTTATTCACAAACCCATCATGAGAGATGGGGTGGCCATATCCCTTAC	795
Dd	232	IleSerGlnGlyLeuPheHisGlyIalIaleMetGluSerGlyValAlaLeuLeuProIy	251
Qy	796	CTGGAGCCCATGATTATGAGAAAGGTAGGACGCTG-----CAGTGGTTCACATTTC	849
Dd	252	LeuIleAla-----SerSerAlaAspValIleSerThrValIalAlaAsnLeu	267
Qy	850	TGTGGTAAACATGCCTCAGACTCTGAGCGCTCTGAGGTGCTGAGGACAAACCCCTCC	909
Dd	268	SerAlaCyAspGlnValAspSerGlnIalaleuValGlyCyLeuAlaArgIyIySerIyS	287
Qy	910	AAGAGCTGTCTGACCTCAGCCAGCCAGAAACAAAGCTTTTCACTCAGATGTTGATGTGCT	969
Dd	288	GluGlnIleLeuAlaIleAsnIyAspProPheIyMetIleProGlyValIalAspIyIal	307
Qy	970	TTCTTTCCTAATGAGCCTCTAGACTCATATTGTTCGAAAGCATTTAAAGCAATCTTCC	1029
Dd	308	PheLeuProIaHisPheProGlnIleuLeuAlaSerAlaAspPheGlnProValProSer	327
Qy	1030	ATCATCGAGTCAATTAACACGAGTGTGGCTTCTCTGCTCT-----ATGAAG	1077
Dd	328	IleValIyGlyValAsnAlaHisGlnIuPheGlyTrpLeuIleProIyValIleMetArgIleIy	347
Qy	1078	GAGGCTCTGAGATCTCAGTGGCTCCAAAGATCCCTTGGCCCTCCATCTGATACAAAC	1137
Dd	348	AspThrGlnIyGlyIleMetAspArgIuIalSerGlnIalAlaLeuGlnIyMetLeuIyThr	367
Qy	1138	ATTCGCAACATCCGCGCTCAGATTTGGACCTTGGGGTAATGATATCTCCATCAAG	1197
Dd	368	LeuIleuMetLeuProIleThrPheGlyAspLeuIyHisGlyIuIyTrIleGlyAspAsn	387
Qy	1198	CACCTCCCTGAGTAAATCCGAGACAGCTTCTTGACCTTGTGAGATGTGTCTTTGTG	1257
Dd	388	GlyAspProGlnIyThrLeuGlnIalGlnIlePheGlnIuMetMetAlaAspSerMetPheIal	407
Qy	1258	GTCCTCGACATGATACAGCTGCATATACAGAGATGCTGGTGCACCTGTCTACTTAT	1317
Dd	408	IleProIalaleuGlnValAlaHisPhe--GlnCySerArgIalProValIyThrIy	426
Qy	1318	GAGTTTGGGACCGGCTCAGTGCCTTTGAAAGCAGAAACCGGCTTTGAAAGCGAC	1377
Dd	427	GluPheGlnHisGlnPheSerTrpLeuIyAsnIleArgProIaHisMetCylAlaAsp	446
Qy	1378	CAGCTGTATGAAGTCCGCTTGTGTCTGGTGGTGCCTTCTGAAAGGAGACATTTGTATG	1437
Dd	447	HisGlyAspGlnLeuProPheIalPhe---ArgSerPhePheGlyGlyAsnTyIleIyS	465
Qy	1438	TTGGAAGAGCCACGAGAGAGAAATTATCTGACCGGAGAAATGATGAAATATCTGGCT	1497

```
Db      466 phe-----ThglugluglulglnleuserArglysmectelystyrrTrprAla 462
QY      1498 ACCTTGTCTGAACCGGGAAATCCTAATAGGGAAGCAACTGTCTTCTGTGCCACCTTAATAAT 1555
         |||||               |
Db       483 AenPhlaIarGaAngllYAenPrIoenglgylguleProHieRrProlLeuPheApr 5020
         |||||
QY      1558 CTGACTTGACGCATTACTTCACCCTGACCTTGACAATGAGCCCTGGACAGAATCATCAAAGA 1611
         |||||
Db       503 GlngluGlunGrnyrlreUeInleubsnneUGnProoalValGIlyArGalaleulySAIA 5222
         |||||
QY      1618 CCGGGGTGTGATTTTTGGACGACGACCATCCCC   1650
         :|||||
Db       523 HisArgLeuGINphetripLysALaleuPro 533

RESULT 15
US-10-102-806-689
; Sequence 689, Application US/10102806
; Publication No. US20030054421A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA103PIC1
CURRENT APPLICATION NUMBER: 2002-03-22
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05881
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/1124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 846
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 689
LENGTH: 549
TYPE: PRt
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (7)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-689

Alignment Scores:
Pred. No.:          5,09e-101           Length:              549
Score:             1230.00            Matches:                248
Percent Similarity:    63.84%        Conservative:           91
Best Local Similarity:    46.70%     Mismatches:              168
Query Match:         38.80%         Indels:                 24
                               Gaps:                   8
DB:                  14

US-10-023-515-3 (1-1746) x US-10-102-806-689 (1-549)
QY      91 GGCGCTCTGTGTCAGAGGGCCAAGAGAACACACAGCTGGATGTGATTCAGGGCAAGCAA 150
         |||
         ::||::|||
Db      16 GlyGlnApeSerAIAserProileargThrthrHisethnglglnVallleuglyserleu 35
         |||||
QY      151 GTCACTGTGTCTGGGAAGCCCTGTGCCTGTGAACGCTTTCCTGTGAACTCCCCTTGTGCT 210
         |||||
         ::||::|||
Db      36 ValHisValLySGlyAlaAsnaLaclYvalGlnThrPheleugLYleProPheaLAlys 55
         |||||
QY      211 CCCCGCGTGGGATCCCTGTGGAATTAAGCAACCCGAGCCGTGATCGCCCTGGATTAAGCTTG 270
         |||||
         |||||
Db      56 ProProleuNGlyProleudnrgrPhealaBProFogUpurofogluserTripserglyVAL 75
         |||||
QY      271 CGAGAAGCACCTCTACCTTAATTGTGTGCTCGCAGAAC-----TCAGAG 315
         |||||
         |||||
         ::||::|||
         ::||::|||
         ::||::|||
         ::||::|||
Db      76 ArgserglyThrtnrhIsProAlemetCybsleunbnabryleuthrALAvalGluSERGLu 95
         |||||
         |||||
QY      316 TGCGTGTCTTAGATCAACACATGCTCAAGGTGATTAACCGGAATTGCGAGTGTCAAAA 375
         ::|||
```

```

Db      96 Pheleu-----SerGlnPheAsnMetThrPheProSerAspSerMetSerGln 111
Qy      376 GACTGCTCTACTGACATCTATGAGCTGCTGCCACGCGCATACAGGCTCCAAAGCTCC 435
Db      112 AspCybSeruYrLeuSerIleYrThrProAlaHisSerHisGluGlySerAsnLeuPro 131
Qy      436 GTCTTGATGTGTGCTCCAGAGAGTCCCTTCAAGACTGGCTCAGCTCCATCTTTGAATGG 495
Db      132 ValMetValTrrIleHisGlyGlyAlaLeuValPheGlyMetAlaSerLeuYrAspGly 151
Qy      496 TCCGCGCTGCTGCTGATGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
Db      152 SerMetLeuAlaAlaLeuGlnAsnValValValIleIleGlyTrpArgLeuGlyVal 171
Qy      556 TTGGTCTTCCACACATGAGATGAGATGCTCGGGGAACTGGAGCCCTTCAAGGACGAG 615
Db      172 LeuGlyPhePheSerThrGlyAspYrHisAlaThrGlyAsnTrpGlyYrLeuAspGln 191
Qy      616 GTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Db      192 ValAlaAlaLeuArgTrpValGlnGlnAsnIleAlaHisPheGlyGlyAsnProAspArg 211
Qy      676 GTGACCATCTTTGGGAGCTCGCGGAGCCATAGTGTCTTACTTACTGCTGCTCC 735
Db      212 ValThrIlePheGlyGlySerAlaGlyThrSerValSerSerLeuValAlaSerPro 231
Qy      736 ATGGCCAAAGCTTATTCACAAAGCCATGAGAGATGGGGTGGCCATCATCCCTTAC 795
Db      232 IleSerGlnGlyLeuPheHisGlyAlaIleMetGluSerGlyValAlaLeuLeuProGly 251
Qy      796 CTGAGAGCCCATGATTATGAGAAAGTGAGACTG-----CAGTGGTTCACACTTC 849
Db      252 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAsnLeu 267
Qy      850 TGTGTGAACAATGCGCTCAGACTCTGAGGCTGCTGAGGCTGCTGAGCAAAACCTCC 909
Db      268 SerAlaCybAspGlnValAspSerGlnAlaLeuValGlyCysLeuArgGlyLysSerLys 287
Qy      910 AAGAGCTGCTGACCTCAGCCAGCAAAACAAAGCTTTCACTCAGATGGTGTGATGCT 969
Db      288 GlnGlnIleLeuAlaIleAsnLysProPheLysMetIleProGlyValValAspGlyVal 307
Qy      970 TTCTTCCATGAGCTCTAGATCTATGCTCTCAGAAAGCATTTAAAGCATTCCTTCC 1029
Db      308 PheLeuProAlaGlnProGlnIleLeuAlaSerAlaAspPheGlnProValProSer 327
Qy      1030 ATCATCGAGTCAATATACACAGATGTGCTCTGCTGCT-----ATGAG 1077
Db      328 IleValGlyValAlaAsnAsnGlnIlePheGlyTrpLeuIleProLysValMetArgIleYr 347
Qy      1078 GAGGCTCTGAGATCTCTCAGTGGCTCCAAAGATCCCTTGCCTCATGTATACAAAC 1137
Db      348 AspThrGlnLysGlnMetCAspArgGlnAlaSerGlnAlaAlaLeuGlnLysMetLeuThr 367
Qy      1138 ATCCGCAACATCCGCTCAGATTTTGCACCTTGGGCTAATGATTAATCTTCATGACAG 1197
Db      368 LeuLeuMetLeuProThrPheGlyAspLeuLeuArgGlnIleYrIleGlyAspAsn 387
Qy      1198 CACTCCCTGACTGAATCCGAGACAGTCTTCTGACTTGTGAGATGTGTTCTTTGTG 1257
Db      388 GlyAspProGlnThrLeuGlnAlaGlnPheGlnGlnIleMetMetAlaAspSerMetPheVal 407
Qy      1258 GTCCCTGCACTGATCAACAGCTCGATATACAGAGATGCTGTGACCTGTCTTACTTAT 1317
Db      408 IleProAlaLeuGlnValAlaHisPhe---GlnCysSerArgAlaProValYrPheYr 426
Qy      1318 GAGTTTGGCAACGGCTCAGAGCTTTGAAGACAGAAAGCGCTTTTGTCAAAGCCGAC 1377
Db      427 GlnPheGlnHisGlnProSerTrpLeuYrAsnIleArgProProHisMetLysAlaAsp 446
Qy      1378 CAGCGTGAATGAAGTCCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
Db      447 HisGlyAspGlnLeuProPheValPhe---ArgSerPhePheGlyGlyLysAntYrIleLys 465

```

```

Qy      1438 TTGGAAGAGCCACGAGAGAGAGAAATTACTGACCCGAGATGATGAATATCTGGCT 1497
Db      466 Phe-----ThrGlnGlnIleGlnIleGlnIleSerArgLysMetMetLysYrTrpAla 482
Qy      1498 ACCTTGCTGCAACCGGATCTTAATGGAAAGACCTGTCTGTGGCCACTTATAT 1557
Db      483 AsnPheAlaArgAsnGlyAsnProAsnGlyGlnIleLeuProHisIleTrpProLeuPheAsp 502
Qy      1558 CTGACTGAGAGTACTCCAGCTGAGCTTGAACATGAGGCTGAGCAGAGACTCAAGAA 1617
Db      503 GlnGlnGlnGlnTrpLeuGlnLeuAsnLeuGlnProAlaValGlyArgAlaLeuYrAla 522
Qy      1618 CCGCGGTGATTTTGGACCAAGCAACATCC 1650
Db      523 HisArgLeuGlnPheTrpLysAlaLeuPro 533

```

Search completed: June 15, 2005, 09:37:11
Job time : 152.806 secs

This Page Blank (uspto)